Processing and presentation of hybrid polypeptides in type 1 diabetes

Harbige, James

Awarding institution:
King's College London

PROCESSING AND PRESENTATION OF HYBRID POLYPEPTIDES IN TYPE 1 DIABETES

BY

JAMES HARBIGE

A thesis submitted for the degree of
Doctor of Philosophy

Department of Immunobiology
School of Immunology & Microbial Sciences
King's College London
May 2020

ABSTRACT

Alleles at the *HLA-DRB1*, *-DQA1* and *-DQB1* loci confer the greatest genetic predisposition to type 1 diabetes (T1D). However, details linking HLA-DQ-mediated predisposition and T cell autoreactivity in T1D are scant. An attractive explanation lies in the discovery of a novel form of modified peptide generated by covalent crosslinking of proinsulin peptides to other pancreatic β-cell peptides, termed hybrid insulin peptides (HIPs). To date, HIPs have been identified in β-cells by mass spectrometry (MS) and HIP-reactive $CD4^+$ T cells detected in the islets and peripheral blood mononuclear cells (PBMCs) of patients with T1D. I hypothesise that the hybrid peptides that have been reported derive from antigen-presenting cell (APC)-dependent processing of an immunogenic hybrid polypeptide. To address this, I optimised an antigen delivery system by conjugating synthetic peptides to carrier proteins, permitting peptide targeting to the APC surface and internalisation. Pulsing of Epstein-Barr virus-transformed B (EBV B) cell lines homozygous for high-risk HLA alleles with a pokeweed mitogen (PWM)-peptide conjugate results in HLA-restricted presentation of a GAD65 epitope and T cell activation. This effect was not observed using peptide-conjugated anti-DEC205 antibody despite evidence of efficient internalisation. Large-scale EBV B cell cultures were pulsed with PWM conjugated to a C-peptide:WE14 hybrid polypeptide and HLA-DR and -DQ peptide complexes isolated for immunoaffinity purification. MS analysis of eluted peptides from HLA-DR4, -DR3, -DQ8 and -DQ2 molecules did not identify hybrid epitopes although evidence of class II processing was observed. Using FluoroSpot to investigate immunogenicity in T1D patient-derived PBMCs, I present data to suggest that immunological processing of hybrid polypeptides to generate a cognate hybrid epitope is limited. Together these findings indicate hybrid epitopes may not be generated through conventional processing and presentation of a long hybrid polypeptide. This project provides novel insight into the mechanism(s) of hybrid peptide generation, highlighting unconventional antigen presentation pathways in T1D.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my first supervisor Professor Mark Peakman to whom I am indebted for allowing me to undertake a PhD project in a topical area of type 1 diabetes research. His mentorship during my studies was invaluable for the project and has undoubtedly shaped my career development as a researcher. I also thank my second supervisor Dr Timothy Tree for his advice and helpful discussion throughout, as well as Dr Michele Mishto who acted as second supervisor toward the end of my PhD. Grateful thanks are due to members of my PhD thesis committee: Professor Peter Jones, Dr Deena Gibbons and Professor Juan Martin-Serrano for their useful suggestions. I am especially indebted to Dr Xiaoping Yang for her expertise in mass spectrometry which has been a pivotal component of this thesis. To all my colleagues within the Peakman and Tree research groups, I wish to express my utmost appreciation for providing both scientific and personal support during the peaks and troughs of my research. I want to draw particular attention to Dr Martin Eichmann. Words cannot describe how instrumental you have been in providing countless stimulating discussions which have bolstered my scientific training. Finally, my deepest gratitude goes to my partner, Michael, and my brothers, Laurence and Jonathan. Without their continued love, support and encouragement this PhD thesis would not have materialised. I dedicate this thesis to my mother and father who, despite raising triplets, managed to stay sane.

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

ABBREVIATIONS

αDEC205	Anti-DEC205 monoclonal antibody
β_2m	Beta-2 microglobulin
7-AAD	7-Aminoactinomycin D
AA	Amino acid
ACN	Acetonitrile
ADCC	Antibody-dependent cellular cytotoxicity
ADS	Antigen delivery system
AEP	Asparaginyl endopeptidase
AIRE	Autoimmune regulator
APC	Antigen-presenting cell
APC	Allophycocyanin
BCR	B cell receptor
BDC	Barbara Davis Center
BLAST	Basic local alignment search tool
BLCL	B lymphoblastoid cell line
bPWM	Biotinylated pokeweed mitogen
BSA	Bovine serum albumin
CBV	Group B coxsackieviruses
CD	Coeliac disease
CDC	Complement-dependent cytotoxicity
cDNA	Complementary deoxyribonucleic acid
CFSE	Carboxyfluorescein succinimidyl ester
ChgA	Chromogranin A
CLIP	Class II-associated invariant chain peptide

CTLA-4	Cytotoxic T-lymphocyte antigen 4
DC	Dendritic cell
ddH_2O	Double-distilled water
DM	Diabetes mellitus
DMP	Dimethyl pimelimidate
DMSO	Dimethyl sulfoxide
DRiPs	Defective ribosomal products
EBV	Epstein-Barr virus
ECACC	European Collection of Authenticated Cell Cultures
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ELISpot	Enzyme-linked immunosorbent spot
ELU	Eluate
ER	Endoplasmic reticulum
ERAPs	Endoplasmic reticulum aminopeptidases
ERC	Endocytic recycling compartment
FA	Formic acid
FACS	Fluorescence-activated cell sorting
FB	FACS buffer
FCS	Foetal calf serum
FDR	False discovery rate
Fmoc	9-Fluorenylmethoxycarbonyl
FOXP3	Forkhead box P3
GAD65	Glutamic acid decarboxylase 65
GFAP	Glial fibrillary acid protein

GWAS	Genome wide association study
HC	Heavy chain
HEL	Hen egg lysozyme
HIP	Hybrid insulin peptide
HIV-1	Human immunodeficiency virus type 1
HLA	Human leukocyte antigen
IA-2	Insulinoma associated antigen 2
IAA	Insulin autoantibodies
IAPP	Islet amyloid polypeptide
IAPP1/2	Islet amyloid polypeptide propeptide 1/2
IFN-γ	Interferon-gamma
IGRP	Islet-specific glucose-6-phosphatase catalytic subunit-related protein
IL-1β	Interleukin-1 beta
IL-10	Interleukin-10
IL-17A	Interleukin-17A
IL-2	Interleukin-2
IL-22	Interleukin-22
IL2RA	Interleukin 2 receptor subunit alpha
ILV	Intraluminal vesicles
INF/Can	Infanrix-hexa/*Candida albicans* antigen
INS-DRiP	Defective ribosomal product of the insulin gene
LC	Light chain
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LEAF	Low endotoxin, azide-free
m/z	Mass-to-charge ratio

MBP	Myelin basic protein
MFI	Median fluorescence intensity
MHC	Major histocompatibility complex
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
mTECs	Medullary thymic epithelial cells
MW	Molecular weight
MWCO	Molecular weight cut off
NOD	Non-obese diabetic
NPPE	Naturally processed and presented epitopes
NPY	Neuropeptide Y
PAS	Protein A sepharose
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PD	Proteome Discoverer
PD-1	Programmed cell death-1
PDL1/2	Programmed death ligand 1/2
PE	Phycoerythrin
PFA	Paraformaldehyde
PHE	Public Health England
pHLA	Peptide-HLA complex
PLP	Proteolipid Protein
PMA	Phorbol myristate acetate
PPI	Preproinsulin
PSM	Peptide-to-spectrum matching

PTM	Post-translational modification
PTPN22	Protein tyrosine phosphatase nonreceptor type 22
PWM	Pokeweed mitogen
RA	Rheumatoid arthritis
REC	Research Ethics Committee
RMS	Root mean squared
RNA	Ribonucleic acid
ROC	Receiver operator characteristic
RP-HPLC	Reverse-phase high performance liquid chromatography
RPMI	Roswell Park Memorial Institute
RV	Resin volume
SB	Strong binder
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SFU	Spot forming unit
SI	Stimulation index
Sulfo-SMCC	Sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate
T1D	Type 1 diabetes
TAP-1/2	Transporter associated with antigen processing-1/2
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween-20
TCR	T cell receptor
TFA	Trifluoroacetic acid
TGF-β	Transforming growth factor beta
Tfh	T follicular helper

TNF-α	Tumour necrosis factor alpha
Treg	Regulatory T cell
ViPR	Virus Pathogen Database and Analysis Resource
WB	Weak binder
ZnT8	Zinc transporter 8

CHAPTER 1: INTRODUCTION

1.1 Type 1 diabetes: an autoimmune disease

Diabetes mellitus (DM) is a chronic metabolic disorder caused by either insulin deficiency or insulin resistance, or both (Kerner et al., 2014). The two most common forms of DM are type 1 and type 2 diabetes (T1D and T2D, respectively). T1D is an autoimmune disease characterised by destruction of insulin-producing β-cells within the pancreas. As a consequence, patients present with persistent hyperglycaemia (fasting blood glucose ≥126 mg/dL, random plasma glucose ≥200 mg/dL, haemoglobin A1C [HbA1c] ≥6.5%) (American Diabetes, 2018) resulting in a high risk of metabolic and vascular complications (Bhattarai et al., 2019). The absolute requirement for exogenous insulin is a defining feature of T1D compared with non-insulin-dependent T2D. Age at diagnosis and the presence of autoantibodies in the peripheral blood against glutamic acid decarboxylase (GAD65), zinc transporter 8 (ZnT8), insulinoma-associated antigen-2 (IA-2) and insulin (IAA) (Couper and Donaghue, 2009; Vermeulen et al., 2011) are also important hallmarks distinguishing T1D from T2D. The appearance of autoantibodies in the serum can predate clinical onset of the disease by months or even years (Krischer et al., 2003; Riley et al., 1990), thus the role of these autoantibodies in disease pathogenesis is still not yet clear.

1.2 Epidemiology

Onset of T1D most commonly occurs in childhood (van Belle et al., 2011) with a peak incidence between 10-14 years of age. However, T1D is no longer considered to be restricted to children and adolescents as the disease can manifest at almost any age from early infancy to late adulthood (Maahs et al., 2010; Rogers et al., 2017; Thunander et al., 2008). T1D is thought to be one of the most common endocrine/metabolic diseases diagnosed in children, representing 80-90% of diabetic children (Craig et al., 2009). The global incidence of T1D is unknown

however, up to 3 million T1D patients were estimated in the USA in 2010 (Chiang et al., 2014), equivalent to 1 in 300 by 18 years of age (Shojaeian and Mehri-Ghahfarrokhi, 2018). Interestingly, wide variations exist between the incidence rates of different populations with China and Venezuela bearing the lowest incidence rates (0.1 per 100 000 per year) and the highest incidence observed in Finland and Sardinia (37 per 100 000 per year) (Soltesz et al., 2007). In general, the incidence rate of T1D is higher in populations in Europe or in populations of European origin (e.g. USA, Canada, Australia and New Zealand) (Diamond Project Group, 2006). Incidence rates in Europe follow a north-south gradient (EURODIAB ACE Study Group, 2000; Green et al., 1992) with Sardinia as an exception, being south of Finland and having a similarly high incidence rate.

This global variation in incidence can be explained in part by variation between populations in the distribution of human leukocyte antigen-DQ (HLA-DQ) genotypes conferring high-risk for T1D (Ronningen et al., 2001). Additionally, differences in environmental risk factors such as nutrition and lifestyle may be an important determinant of incidence rates (Patterson et al., 2001) in combination with genetic susceptibility markers. Worldwide epidemiological data suggests the incidence of T1D is increasing by 2-5% annually (Shojaeian and Mehri-Ghahfarrokhi, 2018) therefore the need to better understand the genetic and environmental factors involved in the pathogenesis continues to grow.

1.3 Aetiology: genetic and environmental risk factors

T1D is a multifactorial disease arising from a complex interaction of genetic and environmental factors. This is evident by the observation that 90-95% of T1D children carry high-risk HLA haplotypes, yet in the general population only 5% or less bearing the haplotype develop T1D (Maahs et al., 2010). Genome-wide association studies (GWAS) have identified >60 loci

associated with risk of developing T1D (Pociot, 2017). However, the strongest association is with genes mapping to the HLA region, specifically those mapping to the HLA class II region, with odds ratios ranging from 0.02 to >11 for specific HLA-DR-DQ haplotypes (Erlich et al., 2008). Class II genes encoding either the HLA-DR4-DQ8 (*HLA-DRB1*04–DQA1*03:01–DQB1*03:02*) or HLA-DR3-DQ2 (*HLA-DRB1*03:01–DQA1*05:01–DQB1*02:01*) haplotypes confer the greatest disease risk (odds ratios = ~5) and the risk is ~5-fold higher again in individuals who are heterozygous for HLA-DR3-DQ2/DR4-DQ8 haplotypes (Pociot and McDermott, 2002). On the contrary, dominant protection from T1D is conferred by possession of the HLA-DQ6 (*HLA-DQA1*01:02–DQB1*06:02*) haplotype. Although associated with lower odds ratios (~2-4), class I HLA genes are also implicated in T1D risk (Nejentsev et al., 2007) with $CD8^+$ T cells shown to kill human β-cells *in vitro* by recognition of HLA-A2 (*HLA-A*02:01*) and HLA-A24 (*HLA-A*24:02*)-restricted preproinsulin epitopes (Knight et al., 2013; Skowera et al., 2008).

Polymorphism in the promoter region of the insulin gene (odds ratio = 2.4) yields the strongest genetic association besides HLA (Pociot et al., 2010) and is thought to result in lower levels of proinsulin gene expression in the thymus, thus shaping negative selection (Nokoff and Rewers, 2013). Only two other loci, protein tyrosine phosphatase nonreceptor type 22 (*PTPN22)* and interleukin 2 receptor alpha (*IL2RA)* have consistently reported odds ratios greater than 1.5 (Pociot, 2017). The latter is associated with lower levels of forkhead box P3 (FOXP3) expression by regulatory T cells (Tregs) and a reduction in their ability to suppress effector T cells (Garg et al., 2012).

It is most likely that T1D arises in genetically susceptible individuals due to an environmental trigger (van Belle et al., 2011). This is strengthened by a discordance of age of T1D

development between monozygotic twins, suggestive of an unshared environmental factor (Redondo et al., 1999). Several environmental factors have been proposed, of which viral infection has gathered the strongest evidence. Enteroviral infection such as group B coxsackieviruses (CBV) have been strongly implicated with one study reporting 67% of children diagnosed with T1D exhibit CBV IgM responses (Friman et al., 1985). Equally, of fourteen serum samples taken from children diagnosed with new-onset T1D, 64% (9/14) showed evidence of enteroviral RNA particularly coxsackie B3 and B4 viruses (Clements et al., 1995). Further evidence was provided by reports suggesting coxsackievirus B1 increases the risk of disease development in children at high-risk or recently diagnosed T1D (Laitinen et al., 2014; Oikarinen et al., 2014). Mouse models have suggested a potentially causative relationship with inoculation of mice with a coxsackievirus (CVB4E2) isolated from a child with diabetic ketoacidosis shown to induce diabetes (Yoon et al., 1979). Furthermore, maternally transferred CVB-specific antibodies in the non-obese diabetic (NOD) mouse reduce the risk of CVB infection in the offspring and protect them from T1D development (Larsson et al., 2013). The exact mechanism by which viral infection contributes to β-cell destruction is yet to be elucidated although molecular mimicry, β-cell dysfunction and bystander activation of T cells have been proposed (Dunne et al., 2019; Xia et al., 2019). Nevertheless, as studies in humans have largely been epidemiological, a conclusive link between viral infection and human T1D development is still lacking.

1.4 T cell mediated immune tolerance

In order to understand autoimmunity and autoimmune disease, it is important to appreciate the mechanisms by which the risk of T cell recognition of self-antigens is, for the most part, held in check. Immature T cells, termed thymocytes, are subject to a selection process in the thymus with the outcome being dependent upon the affinity of the T cell receptor (TCR) for self-

peptide–major histocompatibility complex (MHC) complexes (Hogquist et al., 2005). The essence of thymic self-representation is embodied by the autoimmune regulator transcription factor (AIRE) which permits ectopic expression of tissue-specific antigens (e.g. insulin) by medullary thymic epithelial cells (mTECs) (Anderson et al., 2002). T cells bearing a TCR with high affinity for self-peptide–MHC complexes undergo deletion by negative selection, conversely those TCRs that have no detectable affinity for peptide–MHC ligands are subject to death by neglect. Thymocytes that recognise self-peptide–MHC complexes with low affinity induce positive selection, promoting thymocyte maturation, survival and release into the periphery. Thymic Tregs are generated in response to intermediate-affinity interactions with self-antigen (Klein et al., 2014).

However, not all self-antigens that T cells need to be tolerant to are expressed in the thymus, and thus central tolerance mechanisms are not 100% efficient. The result is the release of potentially autoreactive T cells into the periphery. Peripheral tolerance mechanisms serve to restrain activation and expansion of these self-reactive T cells after their exit from the thymus. These mechanisms include immunological ignorance, i.e. an autoreactive T cell is kept in ignorance by sequestration of the self-antigen behind anatomical barriers such that it never encounters its cognate self-antigen *in vivo*. In addition, encounter of the self-antigen can lead to activation-induced cell death (apoptosis) via the Fas and Fas ligand (FasL) pathway or induce anergy (a state of unresponsiveness), potentially involving interaction of the T cell molecules CTLA-4 or PD-1 with their respective ligands (CD80/86, PD-L1/2) (Walker and Abbas, 2002). In addition, Tregs can directly supress autoreactive T cells in the periphery by producing inhibitory cytokines such as IL-10 and transforming growth factor-β (TGF-β) (O'Garra et al., 2004). A failure to maintain self-tolerance mechanisms results in self-antigens being targeted by the immune system causing chronic inflammation and autoimmunity.

1.5 The dominant role of the adaptive immune system in T1D pathogenesis

Most evidence implicate the adaptive immune system (T and B lymphocytes) as being central to β-cell destruction in T1D (Roep, 2003). That being said, knowledge of the mechanisms by which tolerance to β-cell autoantigens is inadequate or unsuccessful is scant. To date, a model constructed from human and mouse studies, suggests a yet unidentified initiating event results in damage to the islet of Langerhans (Roncarolo and Battaglia, 2007). The inflammation that ensues stimulates the release of β-cell autoantigens which are endocytosed by professional antigen-presenting cells (APCs) e.g. dendritic cells (DCs). Following migration of the APC to the local lymph node, peptides derived from the β-cell autoantigen are presented via HLA class I and II molecules to $CD8^+$ and $CD4^+$ T cells, respectively, culminating in T cell activation.

Functionally impaired Tregs are unable to constrain activation of islet-reactive T cells (Lindley et al., 2005) allowing them to home to the pancreas where $CD8^+$ T cells can directly kill β-cells (Bulek et al., 2012; Kronenberg et al., 2012; Skowera et al., 2008) through release of cytolytic granules containing perforin and granzymes, as well as through Fas–FasL interactions. $CD4^+$ T cells produce pro-inflammatory cytokines such as interferon-γ (IFN-γ) and tumour necrosis factor-α (TNF-α) which promote M1 macrophages resident in the islets to produce reactive oxygen species, TNF-α and IL-1β which in turn augment damage to β-cells (Burrack et al., 2017). Additionally, T follicular helper (Tfh) cells can activate islet antigen-specific B cells, stimulating their differentiation into antibody-producing plasma cells (Serr and Daniel, 2018). Within the inflammatory environment of the islet, activated B cells can function as APCs; antibodies can bind β-cells to mediate antibody dependent cell-mediated cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC), further enhancing the autoimmune response (Wallberg and Cooke, 2013). Due to their ability to concentrate antigen manyfold by virtue of the specificity of the BCR (Wong et al., 2004), B cells have been utilised as APCs to identify

epitopes by elution from HLA class II molecules (Arif et al., 2004; Peakman et al., 1999). However, it is important to appreciate the existence of other APCs such as macrophages and DCs, which may be present at the earlier stages of diabetes (Carrero et al., 2017; Ferris et al., 2014) and may process antigen differently to B cells (van Lummel et al., 2016a).

Increasing evidence exists that innate immune cells also play a critical role in T1D pathogenesis. (Diana et al., 2013) showed in NOD mice that physiological β-cell death induces the recruitment and activation of B-1a cells, neutrophils and plasmacytoid DCs to the pancreas. This infiltration occurs in the first postnatal weeks and innate immune cell crosstalk takes place which is crucial for the development of T1D. A reduction in circulating neutrophils has also been documented in T1D patients and in autoantibody-positive at-risk individuals (Valle et al., 2013). Furthermore, neutrophils are observed at disease onset and at later stages, in the exocrine pancreas of patients with T1D but not in that of T2D or nondiabetic controls. Cytotoxic substances released by neutrophils during maturation, including cytokines, reactive oxygen species and extracellular traps may cause damage to the islet (Huang et al., 2016). Although data on innate immune cells in T1D pathogenesis remains limited, the role of these cells in the initiation and development of T1D will undoubtedly garner further interest in the future.

Importantly, any proposed pathobiological model has to take into consideration that HLA genes at *DRB1*, *DQA1* and *DQB1* loci encode the highest genetic predisposition to T1D (Pociot and McDermott, 2002). A study by (Miyadera et al., 2015) suggested instability of HLA-DQ protein may be important and/or biased binding of specific host or infectious agent-derived antigenic peptides by disease-associated HLA molecules (Nepom and Kwok, 1998). Overall, however, a lack of convincing explanations for HLA-DQ-mediated predisposition exists within the current model.

1.6 The major targets: β-cell autoantigens

This consensus model outlined above suggests that the anti-β-cell immune response is concentrated onto distinct molecular targets for which discovery was largely guided by the recognition of native proteins by islet cell autoantibodies (Roep and Peakman, 2012). Insulin, GAD65, IA-2 and ZnT8 are well established as major autoantigens, due to prognostic and diagnostic features of disease-associated autoantibodies (Achenbach et al., 2009; Ziegler et al., 2013). In recent years, evidence of recognition of several additional antigens has been obtained in human T1D, for example islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP) (Mallone et al., 2007; Yang et al., 2006) which was discovered by reverse translation, chromogranin A (ChgA) (Gottlieb et al., 2014; Li et al., 2015) and islet amyloid polypeptide (IAPP) (Denroche and Verchere, 2018; Panagiotopoulos et al., 2003; Standifer et al., 2006) (Table 1.1). Many of the autoantigens are components of secretory granules which upon release by the β-cell, may facilitate their uptake by APCs and delivery to the HLA class II processing machinery (Pugliese, 2017).

Table 1.1 Established autoantigens in T1D.

Autoantigen	**Cellular expression**	**Subcellular localisation**	**Human T1D**		
			Autoantibodies detected in patients	$CD4^+$ T cell responses	$CD8^+$ T cell responses
Insulin	β-cell	Secretory granule	+	+	+
ZnT8	β-cell	Secretory granule	+	+	+
IA-2	Neuroendocrine	Secretory granule	+	+	+
GAD65	Neuroendocrine	Synaptic-like microvesciles	+	+	+
IGRP	β-cell	Endoplasmic reticulum		+	+
IAPP	β-cell	Secretory granule	+	+	+
ChgA	Neuroendocrine	Secretory granule		+	+

[†]ZnT8, zinc transporter 8; IA-2, insulinoma-associated antigen-2; GAD65, glutamic acid decarboxylase 65; IGRP, islet-specific glucose-6-phosphatase catalytic subunit-related protein; IAPP, islet amyloid polypeptide; ChgA, chromogranin A. Adapted from (Purcell et al., 2019b).

1.7 Classical antigen processing and presentation

In canonical antigen presentation via MHC class I, antigenic peptides presented to $CD8^+$ T cells are generated in the cytosol by a multi-catalytic complex called the proteasome (Groettrup et al., 2010). The proteasome degrades cellular proteins into peptide fragments. The peptides are then transported into the endoplasmic reticulum (ER) lumen by the transporter associated with antigen processing 1 and 2 (TAP-1 and TAP-2). In the ER, these peptides are further trimmed to approximately 8-10 amino acids in length by ER aminopeptidases (ERAPs). With the help of the peptide loading complex, peptides are loaded onto the MHC class I heavy chain-β_2m complex in the ER and the peptide–MHC complex is transported to the cell surface.

Conversely, exogenous antigens destined for presentation by HLA class II molecules enter the endocytic pathway of APCs by multiple mechanisms (Roche and Furuta, 2015). An example of one such process is receptor-mediated endocytosis via clarithin-coated pits which requires antigen binding to receptors present on the APC surface, e.g. lectin receptors, resulting in internalisation of the antigen-receptor complex and delivery into early endosomes. Macropinocytosis represents another mechanism, dependent on the actin cytoskeleton to mediate the non-selective uptake of solute molecules giving rise to large endocytic vacuoles called macropinosomes (Lim and Gleeson, 2011). Arguably the most important mechanism of antigen uptake is phagocytosis, whereby opsonised antigens bind to various cell surface receptors to stimulate re-organisation of the actin cytoskeleton surrounding the antigen (Stuart and Ezekowitz, 2005). By this process, antigens are delivered into the cell in membrane-derived phagosomes. In each of the aforementioned processes, the internalised early endosomes fuse with late endosomal–lysosomal compartments rich in proteolytic enzymes that are active in the low pH environment maintained within this compartment for antigen proteolysis (Blum et al., 2013). It is in these late endosomal–lysosomal compartments where peptide–MHC class II

complex formation occurs. The process by which membranes encapsulate cytosolic antigens to form an autophagosome, termed autophagy, is also an important mechanism of antigen acquisition (Schmid et al., 2007). In a similar manner to the other mechanisms of antigen uptake, the autophagosome eventually fuses with lysosomal compartments to form an autophagolysosome where peptide–MHC class II complexes are generated.

In the ER, newly synthesised MHC class II αβ dimers associate with the invariant chain (Ii), transit through the Golgi apparatus and are targeted to the plasma membrane by signal sequences present in the Ii chain (Figure 1.1) (Cresswell, 1996). At the cell surface, Ii:MHC class II complexes undergo clathrin-mediated internalisation into early endosomes which eventually fuse with late endosomal–lysosomal compartments containing antigenic proteins and peptides. Proteolytic cleavage of the Ii chain generates a fragment of Ii, termed class II-associated invariant chain peptide (CLIP), which remains bound in the peptide binding groove of the MHC class II molecule. Removal of CLIP is mediated by the enzyme HLA-DM allowing the nascent MHC class II molecule to bind antigenic peptides. The activity of HLA-DM is regulated by HLA-DO (Denzin et al., 2005). Peptide–MHC class II complexes are transported via vesicles which deposit peptide–MHC class II directly into the membrane for $CD4^+$ T cell recognition. If recognition by $CD4^+$ T cells does not occur, surface peptide–MHC class II complexes can be internalised by clathrin-independent endocytosis and are targeted for lysosomal degradation or via recycling endosomes back to the plasma membrane (Cho and Roche, 2013).

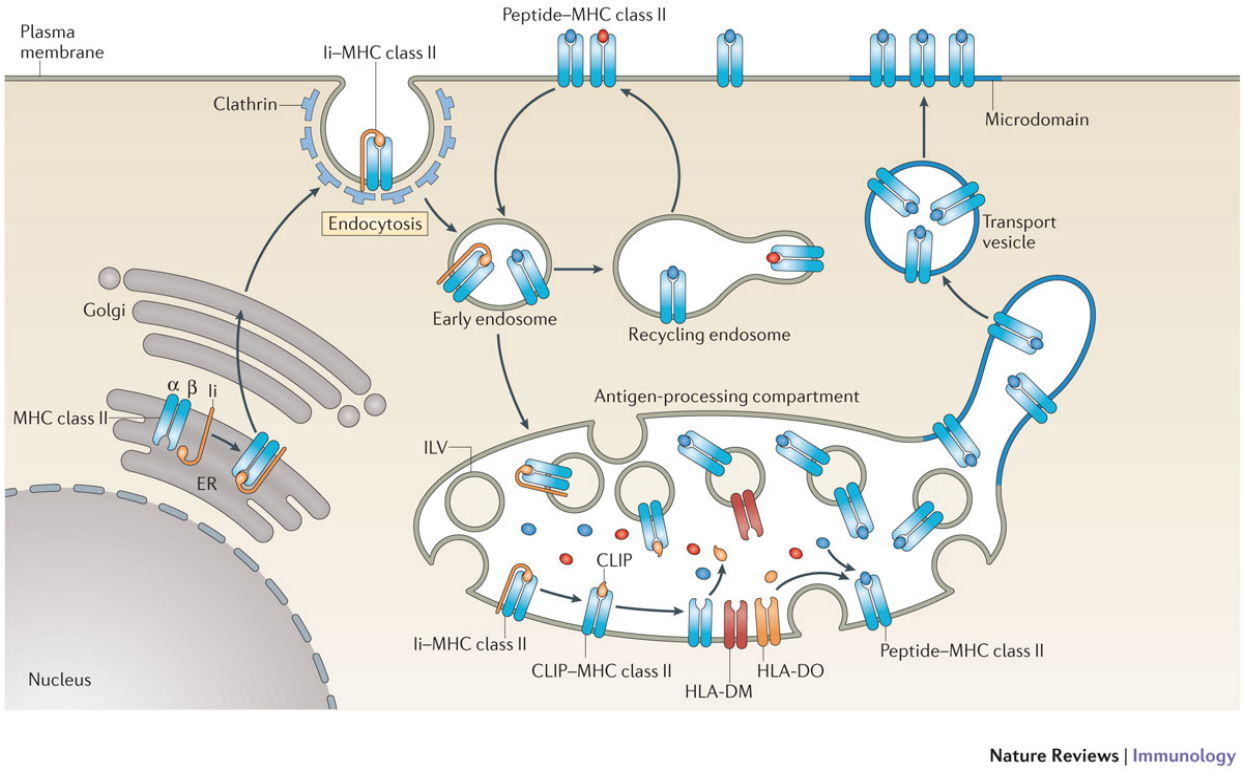


Figure 1.1 MHC class II presentation pathway. ER, endoplasmic reticulum; ILV, intraluminal vesicles; Ii, invariant chain. Taken from (Roche and Furuta, 2015).

1.8 Non-conventional epitope generation

The inadequacy of the consensus model – the lack of a convincing explanation for HLA-DQ-mediated susceptibility – has fostered novel insights into the mechanisms of loss of β-cell tolerance. The discovery of neo-epitopes, also termed non-conventional epitopes (an epitope that is derived after modification of the germline sequence), as T cell targets in T1D has gathered much attention in recent years and may account for the disease association with specific HLA alleles.

1.8.1 Post-translational modification of autoantigens

Post-translational modification (PTM) of self-proteins could generate neo-antigens which can engage actively with the peripheral immune system due to insufficient thymic education. Thus, PTM of antigens has been extensively studied in autoimmune/chronic inflammatory diseases

such as rheumatoid arthritis (RA) (Trouw et al., 2017) and coeliac disease (CD) (Koning et al., 2015; Sollid and Jabri, 2011). A list of common PTMs associated with autoimmune diseases is provided in Table 1.2. Of particular importance, is the enhanced MHC binding capacity described for many post-translationally modified peptides. In RA, citrullination of aggrecan and vimentin epitopes enhances binding of these peptides to disease-associated HLA-DR4 molecules (Scally et al., 2013). Equally, deamidation of gluten-derived peptides by tissue transglutaminase in CD increases binding affinity to disease-associated HLA-DQ2 and -DQ8 molecules, increasing their dwell-time and as a consequence results in a strong T cell response (Molberg et al., 1998; van de Wal et al., 1998). T1D shares similar HLA associations with RA and CD, thus it is plausible that autoantigen modification may have a similar role in T1D pathogenesis.

Table 1.2 Common post-translational modifications of self-antigens associated with autoimmune disease.

Autoimmune disease	Modification	Antigen	Reference
Multiple sclerosis	Citrullination	MBP, GFAP	(Bradford et al., 2014)
	Acetylation	MBP	(Zamvil et al., 1986)
	Palmitoylation	PLP	(Greer et al., 2001)
Rheumatoid arthritis	Citrullination	Vimentin, aggrecan	(Scally et al., 2013)
	Deimination	Fibrin	(Masson-Bessiere et al., 2001)
Systemic lupus erythematosus	Phosphorylation	RNA splicing factors	(Neugebauer et al., 2000)
	Deimination	Histones	(Dwivedi et al., 2014)
Celiac disease	Deamidation	Gluten (gliadin)	(Arentz-Hansen et al., 2000)
Psoriasis	Endoprotease cleavage	Pso27	(Iversen et al., 2011)
Type 1 diabetes	Deamidation	Proinsulin	(van Lummel et al., 2014)
	Disulfide bond	Insulin	(Mannering et al., 2005)
	Citrullination	GAD65	(McGinty et al., 2014)
	Oxidation	Insulin	(Strollo et al., 2017)

[†]MBP, myelin basic protein; GFAP, glial fibrillary acidic protein; PLP, myelin proteolipid protein; GAD65, glutamic acid decarboxylase 65. Taken from (Harbige et al., 2017).

T cell responses to a disulphide-modified insulin A chain epitope (A1-13) restricted by HLA-DR4, was the first evidence that post-translationally modified peptides could be relevant in T1D (Mannering et al., 2005). Since then additional evidence has come from the identification of a transglutaminated ChgA peptide which exhibited increased antigenicity to NOD mouse-derived diabetogenic $CD4^+$ T cell clones (Delong et al., 2012). Furthermore, deamidation of islet peptides by the tissue transglutaminase enzyme enhances their binding affinity to HLA-DQ8 and HLA-DQ8trans molecules (van Lummel et al., 2014). Isolation of a T cell clone from a new-onset T1D patient capable of recognising a deamidated proinsulin peptide presented by HLA-DQ8 provides *a priori* evidence that such modifications could be relevant *in vivo*. HLA-DR4-restricted citrullinated and transglutaminated GAD65 epitopes have also been described with T cells reactive against these modified epitopes present at significantly higher frequencies in T1D patients than HLA-matched controls (McGinty et al., 2014). Altogether the data indicate that immune recognition of enzymatically-modified epitopes does occur in T1D although the link to disease pathogenesis is yet to be uncovered.

Unlike RA and CD, the PTM model in T1D is yet to be firmly consolidated into our understanding of disease pathogenesis. The field has therefore sought new insights into how non-conventional epitopes are generated and their relevance to the conventional peptide repertoire. The discovery of hybrid insulin peptides (HIPs) as a novel form of modified peptide containing non-germline encoded epitope sequences, inspired new thinking around the mechanisms of epitope generation in T1D (Delong et al., 2016). HIPs represent peptides formed not by modification of amino acids (discussed above) but fusion of amino acids sequences. Although in the context of T1D several reports implicate this more complex modification to be involved in the generation of MHC class II epitopes, it has also been reported to give rise to MHC class I-restricted peptides (spliced peptides) (Faridi et al., 2018; Liepe et al., 2016). To date only one report of a HLA class I-restricted IAPP fusion peptide with potential relevance

to T1D has been demonstrated (Gonzalez-Duque et al., 2018). Yet another example which brought non-conventional epitopes into the limelight was the discovery of a defective ribosomal insulin gene product that is targeted by T cells in T1D patients (Kracht et al., 2017). Low or absent expression of these non-conventional epitopes in the thymus may bypass central tolerance and in certain circumstances may be able to explain the highly predisposing *HLA-DQ* gene effect.

1.8.2 Hybrid insulin peptides

First identified in the NOD mouse (Stadinski et al., 2010a) and later in T1D patients (Gottlieb et al., 2014), the natural ChgA-processed peptide, WE14, is a weak agonist for several ChgA-reactive T cell clones derived from NOD mice including the archetypal Barbara Davis Centre (BDC)-2.5 clone (Stadinski et al., 2010a). By modelling the interaction of the WE14 peptide with IA-g7 (NOD mouse MHC II molecule) the authors were able to show that WE14 occupies only half of the IA-g7 peptide binding groove. Thus, it was proposed that the natural ligand recognised by these islet antigen-reactive $CD4^+$ T cell clones *in vivo* may be a modified form of the WE14 peptide. In line with this hypothesis, (Jin et al., 2015) described that addition of RLGL to the N-terminus of WE14 (**RLGL**WSRMDQLAKELTAE) places the WSRMD motif in the correct position in the IA-g7 peptide binding groove. The N-terminally extended WE14 peptide was demonstrated to be much more potent than unmodified WE14 in T cell stimulation and activates a diverse set of ChgA-reactive T cell clones. Later, a hybrid peptide formed by the fusion of a proinsulin C-peptide fragment on the N-terminal side to WE14 on the C-terminal side, significantly increases activation of the BDC-2.5 T cell clone compared to native WE14 (Delong et al., 2016). It was this discovery which brought HIP formation to the forefront of T1D research, generating an unconventional type of neo-epitope potentially of great relevance to disease pathogenesis.

Currently, HIPs that have been described originate from the covalent crosslinking of proinsulin C-peptide to a peptide from another β-cell secretory granule protein. It is therefore reasonable to surmise that HIP formation occurs preferentially in a protein-dense environment such as within the β-cell granule, currently estimated to contain 50–100 distinct proteins (Brunner et al., 2007; Hickey et al., 2009; Schvartz et al., 2012) many of which have unknown functions (Suckale and Solimena, 2010). The identification of hybrid peptides containing different peptide species, but all derivates of proteins present within the granule (Delong et al., 2016) would support this notion.

Initially, hybrid peptides formed by the fusion of C-peptide with WE14 and IAPP2 (islet amyloid polypeptide propeptide 2) were found to elicit an IFN-γ response using a set of NOD mouse-derived BDC T cell clones and this response was stronger than that of the unmodified epitope. Up until now BDC T cell clones have been reported to have defined reactivities, with these findings now suggesting that HIPs may represent the natural ligands for these clones. To show that HIPs exist *in vivo* within β-cells, mass spectrometric analysis of the granule fraction of insulinoma cells revealed the presence of a HIP containing the DLQTLAL C-peptide sequence fused to the WSRM WE14 sequence (Delong et al., 2016). T cells reactive against these HIPs were consequently isolated from the residual islets of deceased donors with T1D. In this case, the islet-derived T cell clones responded to HIPs formed by the fusion of C-peptide to IAPP2 and neuropeptide Y (NPY) but native C-peptide stimulated these T cells weakly. Intriguingly, these T cells only recognised HIPs presented by APCs in the context of HLA-DQ8, suggesting HIPs may provide the missing link for HLA-DQ-mediated susceptibility. Two hybrid peptides (C-peptide:IAPP2 and insulin-2 C-peptide:insulin-1 C-peptide) were later independently identified in the granule fraction of NOD mouse β-cells providing further confirmation of the existence of HIPs *in vivo* (Wan et al., 2018).

In a study by (Babon et al., 2016), $CD4^+$ T cell lines expanded from islets of cadaveric donors with T1D were shown to respond to some of the same hybrid epitopes described by (Delong et al., 2016) by production of IFN-γ. T cell reactivity was observed to HIPs containing C-peptide fused with IAPP2, IAPP1 (islet amyloid polypeptide propeptide 1) and the A chain of insulin. Whether or not these islet-derived T cell lines cross-react with the unmodified peptide and whether the response to the HIP is different to that of the unmodified version was not elucidated. The capacity of C-peptide:IAPP HIPs to stimulate the IAPP-specific BDC-6.9 T cell clone was shown by (Wiles et al., 2017) to be dependent on residues from both insulin and IAPP components, since N- or C- terminal truncation abolished IFN-γ secretion by the BDC-6.9 clone. In addition, modification of the hybrid sequence revealed that negatively charged amino acid side chains are preferred at the C-terminus and that this may be necessary for optimal T cell activation. Similar to HLA-DQ8 molecules, IA-g7 preferentially binds peptides with C-terminal negatively charged amino acids (Suri et al., 2005). Thus, hybrid peptides containing negatively charged amino acids at anchor positions may provide optimal binding to HLA molecules.

(Baker et al., 2018) advanced our knowledge of HIPs in T1D by using IA-g7 tetramers loaded with HIPs to track HIP-reactive T cells over the progression of the disease. Both nondiabetic and diabetic NOD mice harbour high frequencies of T cells reactive to a C-peptide:WE14 hybrid peptide in the pancreas, with T cells reactive to a C-peptide:IAPP2 HIP also being present in the pancreas of NOD mice. An important finding was that antigen-experienced HIP-reactive T cells were detected in the pancreatic lymph nodes well in advance of their infiltration the pancreas (as early as 3 weeks of age). This suggests that HIP-reactive T cells acquire this phenotype early on in the disease process and are amongst the first antigen-specific $CD4^+$ T cells to be activated in the pancreatic lymph nodes. After shedding L-selectin (CD62L) and exit

from the lymph nodes, they can be detected in peripheral blood with increasing frequency throughout disease progression. In this manner, T cells reactive to HIPs may be useful biomarkers of disease and it would be curious to determine if this is also true in humans.

Interestingly, very few HIP-reactive T cells express FOXP3 in the pancreas at the time of disease onset (Baker et al., 2018). Conversely, a large frequency (>20%) of insulin B9-23-reactive T cells in the pancreas expressed FOXP3 indicating that these cells might be involved in immune regulation. These findings were later confirmed by (Ito et al., 2018) and (Spence et al., 2018) who reported that islet-infiltrating Tregs express TCRs specific for B9-23 but not for HIPs. The former study also observed that HIP-reactive T cells were skewed towards a Th1 phenotype upon antigen encounter, secreting both TNF-α and IFN-γ (Ito et al., 2018). Overall, these studies suggest differences, which may lie at the level of thymic education, in $CD4^+$ T cell regulation to conventional and non-conventional epitopes. Restoring a lack of regulatory function in HIP-reactive T cells could therefore be useful in immunotherapeutic approaches.

As research on HIPs in T1D has gathered momentum, an important question is whether individuals that do not develop diabetes also form HIPs. Criteria for the accurate identification of these peptides by mass spectrometry was developed (Wiles et al., 2019) and the presence of HIPs confirmed in primary NOD mouse and human islets. Surprisingly, HIPs were also identified in islets from non-diabetes prone BALB/c mice and islets from human donors without T1D. This indicates that the presence of HIPs alone is not sufficient to cause disease. It is likely that in predisposed individuals, differences in the immune response could precipitate disease development. Such differences include how efficiently HIPs are presented by HLA molecules, how efficient tolerance to HIPs is established and the degree to which they are generated (Baker et al., 2019a; Wiles et al., 2019).

Evidence for the existence of HIPs in human islets has become an increasingly important research focus. Further to the presence of HIP-reactive T cells in the islets of T1D patients, HIP-reactive T cells can also be detected in the peripheral blood of newly diagnosed patients (Baker et al., 2019b). Approximately 30% of T1D patients elicit an IFN-γ response to at least one HIP with responses to up to seven HIPs detected in some patients. In contrast, in HLA-matched controls T cells responses to HIPs were less frequent, of lower magnitude and seldom to more than one HIP. Longitudinal data from one patient indicated that T cell responses to HIPs can continue up to 1 year after diagnosis. Subsequent isolation of T cell clones with different HIP-specificities showed reactivity to target HIPs at low nanomolar concentrations but not to native epitopes, confirming the findings of (Delong et al., 2016). One T cell clone was isolated from the same patient at two different time points, indicating that HIP-reactive T cells traffic in the blood for a prolonged period after onset of T1D. Given that most C-peptide-specific T cell clones derived from the peripheral blood of T1D patients are HLA-DQ-restricted (So et al., 2018), it would be of interest to investigate whether these clones also cross-react with putative HIPs recently identified in human islets (Wiles et al., 2019). Overall, the unearthing of HIPs as T cell targets in T1D has questioned current concepts of disease pathogenesis.

There are, however, limitations which need to be resolved before non-conventional epitopes such as HIPs can be accommodated into our dogma of disease development. Although HIPs have been identified *in vivo* in human islets (Wiles et al., 2019), whether naturally processed and presented HIP epitopes are generated via HLA class II (and class I) processing pathways is undetermined. Furthermore, their relationship to the natural peptide repertoire is also unknown. For example, it is possible that an early autoimmune response is primed against conventional epitopes of C-peptide and/or WE14 but this is insufficient for autoimmunity to be initiated. Encounter with a non-conventional epitope such as a HIP subsequently acts as a super-agonist,

stimulating expansion of diabetogenic T cell clones which drive disease (Figure 1.2). In this two-step model, a breakdown of tolerance would precede an event (e.g. local inflammation) which results in the formation of non-conventional epitopes. Conversely, the initial autoreactive response is primed to non-conventional epitopes *de novo* (tolerance breaker). Such activated T cells may cross-react with germline epitopes, opening up further opportunities for TCR engagement. The importance of higher HLA binding and/or TCR recognition of HIPs compared with germline epitopes is also an unknown at this stage as is the initial cue responsible for their generation. Such an event may arise under different circumstances such as viral infection or weaning where the islet undergoes significant β-cell re-modelling (Gunasekaran et al., 2012) and may be augmented by inflammation and/or ER stress.

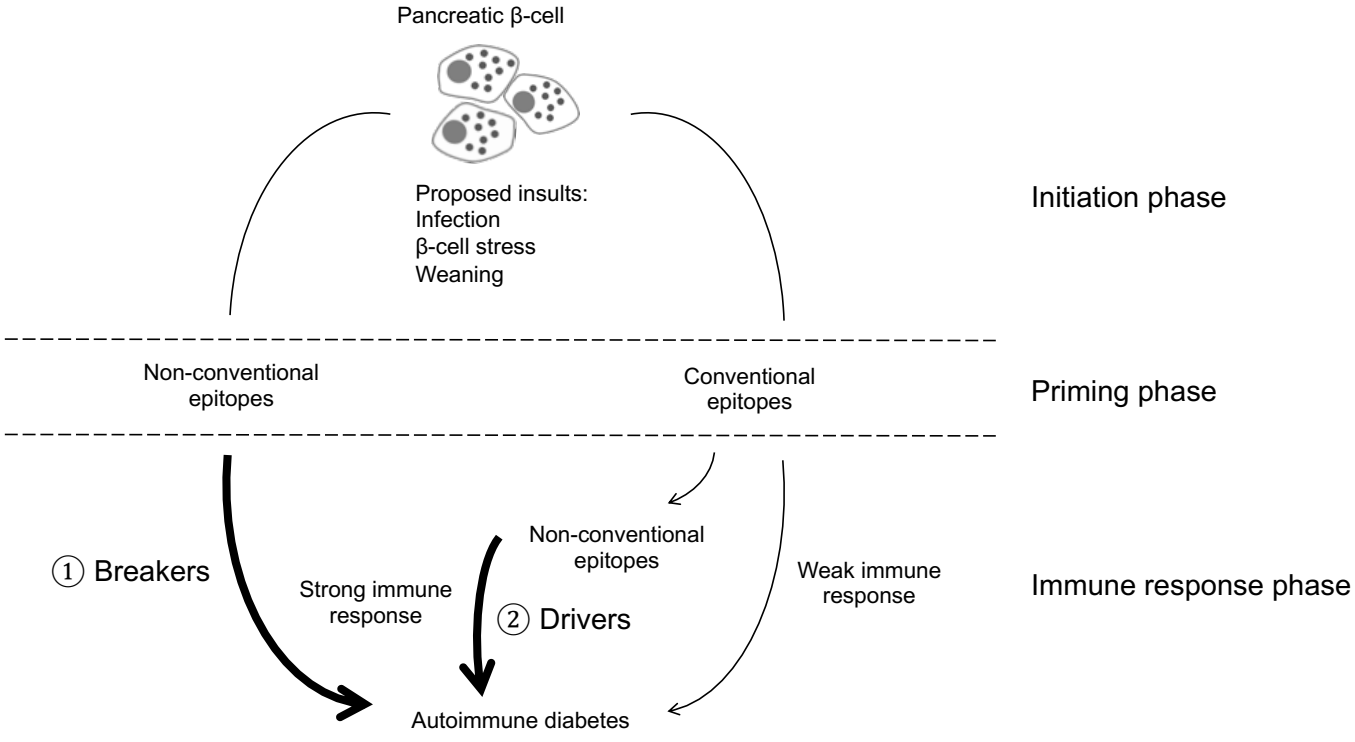


Figure 1.2 Non-conventional epitopes: tolerance breakers (①) or disease drivers (②) in autoimmune diabetes. The initial autoreactive T cell response is primed to non-conventional epitopes such as HIPs (tolerance breaker). Alternatively, non-conventional epitopes complement a pre-existing autoimmune response to conventional epitopes to drive disease. HIPs, hybrid insulin peptide. Taken from (Harbige et al., 2017).

1.8.3 Spliced peptides

Peptide splicing warrants extensive discussion here as it is a relatively novel extension of the conventional view of processing through the canonical HLA class I pathway. Since its first discovery, peptide splicing has been shown to be relevant to cancer pathology and infectious disease (Mishto and Liepe, 2017). For these reasons, it is illustrative of some of the important questions surrounding hybrid peptides in T1D, i.e. mechanism of generation and the complexity of the processes (requiring computational power because of all of the combinatorial possibilities).

Spliced peptides represent an additional domain of non-conventional epitope generation whereby a non-germline sequence is formed by the fusion of two peptide fragments. Such peptides can originate from splicing of two distinct regions within the same protein (*cis*-spliced peptides) or the fusion of peptide fragments from two distinct proteins (*trans*-spliced peptides). Both mechanisms of spliced peptide generation have been described to be operative in the generation of novel class I epitopes. Furthermore, spliced peptides have been shown to be the result of a post-translational event (Hanada et al., 2004)**;** either fusing two peptide fragments end-to-end (Vigneron et al., 2004) or in reverse order (Michaux et al., 2014; Warren et al., 2006).

By generating a database restricted to peptides generated by *cis*-splicing, (Liepe et al., 2016) showed that *cis*-splicing of peptides, specifically proteasome-catalysed peptide splicing is not scarce within the human immunopeptidome. Analysis of the HLA class I immunopeptidome of cell lines and primary human fibroblasts by mass spectrometry revealed that up to one-third of uniquely identified epitopes are derivatives of spliced species. Interestingly, spliced epitopes constituent one-fourth of the total peptide repertoire in terms of abundance, indicating

comparable efficiency in the generation of non-spliced and spliced epitopes. Indeed, comparable abundance of spliced and non-spliced epitopes has also been suggested by (Ebstein et al., 2016) using $CD8^+$ T cell clones specific for melanoma antigens. Spliced epitopes should therefore be equally likely to trigger $CD8^+$ T cell responses as non-spliced epitopes. Intriguingly, (Liepe et al., 2016) also showed that some proteins were exclusively represented by spliced peptides and not non-spliced peptides. These proteins may not possess linear peptides with sufficient HLA binding capacity and/or generate smaller peptides, and thus splicing is a way to generate peptides of sufficient length and motif for HLA class I loading (Vigneron et al., 2017).

Until recently, *trans*-splicing was predicted to occur at a low frequency simply because two distinct protein substrates can hardly fit inside the proteasome simultaneously, coupled with the low chances of having two protein substrates repeatedly degraded at the same time within the proteasome. As a result, only a handful of HLA class I *trans*-spliced peptides had been shown to occur *in vitro* and *in vivo* (Dalet et al., 2010) but due to increases in computational power, (Faridi et al., 2018) were able to report a considerable contribution of *trans*-spliced peptides to the immunopeptidome. The unexpected proportion of *trans*-spliced peptides highlights additional complexity of the immunopeptidome beyond that previously described by (Liepe et al., 2016). The authors propose that for *trans*-splicing it is not necessary that each of the two peptide fragments consistently originate from a particular region of a protein. In the case of short fragments of spliced peptides, the proteasome could ligate identical polypeptides generated from a multitude of donor proteins. They therefore argue that despite the low probability of an individual *trans*-splicing reaction, a high abundance of "*trans*-donors" could increase the probability.

Peptide joining does not appear to be a random process but represents a finely tuned mechanism which has increased efficiency when the splice reactants bear specific properties (Berkers et al., 2015). For example, the C-terminal splice reactant involved in the generation of one specific spliced epitope was required to be at least three amino acids in length (Michaux et al., 2014). The N-terminal splice reactant is also likely to have a minimal length although N-terminally elongated spliced peptides can be trimmed in the ER by ERAPs (Vigneron et al., 2017). Furthermore, another spliced epitope was shown to be generated with reduced efficiency as the length of the intervening sequence (the sequence excised between two splice-reactants) increased (Dalet et al., 2010). However, studies involving analysis of the entire HLA class I immunopeptidome did not identify a correlation between the frequency at which the spliced peptide is generated and the intervening sequence length (Liepe et al., 2016). In breast and colon cancer cell lines, spliced peptides make up approximately 20% of the class I immunopeptidome with source antigens being preferentially long, hydrophobic and basic (Liepe et al., 2019). An explanation as to why these antigens possess such properties is not yet clear but it is possible to speculate that short antigens are less able to generate spliced peptides simply due to less combinational possibilities.

The implication from the above studies is that peptide splicing is subject to regulation. Precisely how *cis*- and *trans*-spliced peptides are generated and how they contribute to T cell selection in unknown. Both standard and the immunoproteasome have the capacity to splice peptides (Dalet et al., 2011) as does the yeast proteasome (Mishto et al., 2012). Thus, because spliced peptides are by-products of proteasomal activity, any type of proteasome including the thymoproteasome should be capable of generating spliced peptides (Vigneron et al., 2019). Therefore, spliced peptides may not be intrinsically more immunogenic than non-spliced peptides as central tolerance mechanisms may exist. Indeed, a large proportion of spliced

peptides are presented at the cell surface and are immunologically silent (Vigneron et al., 2019). Furthermore, (Liepe et al., 2019) and (Faridi et al., 2018) show overall similarity in the sequence motifs (amino acid preference and anchor residues) of spliced and non-spliced epitopes. This is unsurprising as spliced peptides obey the same antigen processing pathway as non-spliced peptides and are selected based on their affinity for the class I peptide binding groove downstream of the proteasome. Where spliced epitopes may differ to non-spliced epitopes is when spliced peptides escape central tolerance, for example because the source protein is not adequately expressed in the thymus or because the proteasome isoform is inefficient in generating these peptides.

Undoubtedly, the combined actions of *cis*- and *trans*-spliced peptides enable a greater breath of antigen-derived epitopes to be displayed for recognition by T cells, especially important during bacterial or viral infection (Faridi et al., 2018; Liepe et al., 2016). Indeed, there is evidence for the immunogenicity of pathogen-derived spliced epitopes during *Listeria monocytogenes* infection (Platteel et al., 2016; Platteel et al., 2017a). Splicing may also be particularly important for the host upon infection with pathogens with small genomes. Such genomes may not encode adequate class I ligands, thus peptide ligation with other antigens may enhance immunosurveillance. More recently, peptide splicing has been shown to have a role in limiting viral escape and control of HIV-1 replication (Paes et al., 2019). Furthermore, T cell responses against spliced epitopes are documented in cancer, with melanoma gp100 ($gp100^{mel}$)-derived spliced epitopes triggering activation of $CD8^+$ T cells found in the peripheral blood of half of melanoma patients (Ebstein et al., 2016). It is not therefore surprising that *in silico-in vitro* pipelines are being developed to identify tumour-specific spliced epitope candidates which can be subsequently validated as putative targets for anti-cancer therapies (Mishto et al., 2019).

To summarise, data on the relevance of *cis*- and *trans*-spliced peptides in immunopeptidomes on autoimmunity and T1D in particular is limited, with this being an area of discovery in the future. Recently, a HLA-A2-restricted fusion epitope was found to be naturally processed and presented by β-cells, formed via the fusion of $IAPP_{15-17}$ with $IAPP_{5-10}$ (Gonzalez-Duque et al., 2018). However, a pathogenic role of the $CD8^+$ T cells recognising this epitope remains to be definitively established. Based on this, it is not unreasonable to speculate that in an inflammatory milieu, ER stress in β-cells may trigger an upregulation of proteasome-mediated protein degradation which could augment spliced peptide generation.

1.8.4 Generation of hybrid and spliced peptides

The mechanism by which hybrid peptides are generated *in vivo* is not fully understood although it is currently thought to involve transpeptidation (Delong et al., 2016), namely the process by which two peptide fragments are ligated together to form a hybrid peptide (Berkers et al., 2009). The secretory granule represents a candidate organelle for this reaction to occur due to high concentrations of enzymes and β-cell proteins such as insulin (Rorsman and Renstrom, 2003), ChgA and IAPP, which promote the transpeptidation reaction (Delong et al., 2016). The degree of enzymatic involvement, or whether this process occurs spontaneously within the β-cell granule due to the sheer abundance of secretory proteins (molecular crowding) is unknown at present. It is possible to envisage that if an enzyme responsible for HIP formation did exist, it would likely reside in the secretory compartment where MHC class II peptides are produced (Vigneron et al., 2017). The proteasome has been shown to catalyse transpeptidation during the production of epitopes for HLA class I presentation (see below), therefore there is the prospect that hybrid peptides may be proteasome-generated particularly during ER stress or proinsulin misfolding.

The mechanism by which class I spliced peptides are generated has been well-established as a proteasome-dependent mechanism. Until a few years ago, only a handful of human spliced epitopes derived from tumour-associated antigens had been reported (Dalet et al., 2011; Hanada et al., 2004; Michaux et al., 2014; Vigneron et al., 2004; Warren et al., 2006), thus tagging proteasome-catalysed peptide splicing as a rare event. The mechanism, however, has been recognised to involve a transpeptidation reaction (Dalet et al., 2010; Mishto et al., 2012; Vigneron et al., 2004; Warren et al., 2006). The reaction begins with breakage of the peptide bond between a residue (termed P_1) of the N-terminal splice reactant and the succeeding intervening sequence (Figure 1.3). This generates an acyl-enzyme intermediate between the P_1 residue and a catalytic residue of the proteasome facilitating the release of the intervening sequence. In conventional antigen processing, the acyl-enzyme intermediate reacts with a water molecule (hydrolysis) to release the N-terminal peptide fragment. During transpeptidation, the free amino group of a C-terminal splice reactant reacts with the acyl-enzyme intermediate resulting in the joining of the N- and C-terminal splice reactants. Within the proteasome's catalytic pockets, it has been suggested that substrate-specific binding sites exist which mediate peptide splicing (Michaux et al., 2014) but it is not known whether the same binding sites are utilised for conventional antigen processing.

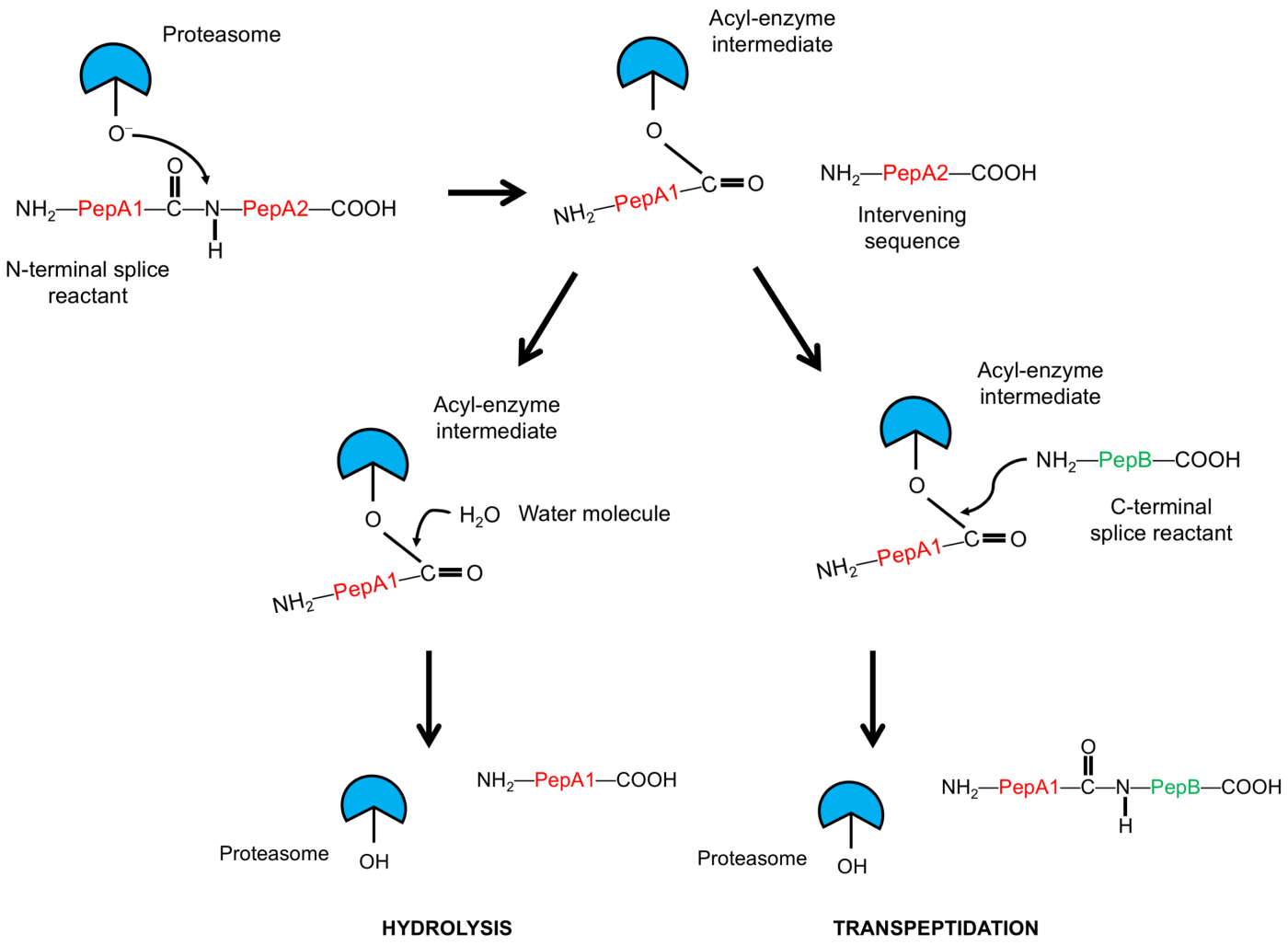


Figure 1.3 Schematic mechanism of proteasome-catalysed peptide splicing by transpeptidation. Adapted from (Jin et al., 2015).

Due to bioinformatic approaches and novel database search strategies, (Liepe et al., 2016) were able to identify thousands of *cis*-spliced peptides in the HLA I immunopeptidome and (Faridi et al., 2018) later showed that *trans*-spliced peptides are as frequently identified as *cis*-spliced peptides found bound to a number of different HLA-A and HLA-B alleles. Thus, the generation of spliced peptides represents a key source of antigenic epitopes. This phenomenon of proteasome-catalysed peptide splicing has to date only been described for class I peptides. Peptide splicing can also give rise to class II epitopes as is the case for hybrid peptides in T1D. These peptides are identified by the ligation of peptide fragments originating from two different secretory granule proteins (by definition *trans*-splicing). The T1D field should proceed with

caution in mixing those class II results with the occurrence of spliced peptides in class I. The generation of *trans*-spliced class II peptides involves a different pathway and may involve different mechanisms to proteasome-mediated peptide splicing of class I epitopes e.g. endosome- or lysosome-based (Harbige et al., 2017). It is possible to envisage that if a pool of spliced species exists in the cytoplasm for HLA class I loading, β-cell death may facilitate their uptake by APCs delivering them into the HLA class II processing and presentation pathway. Alternatively, extracellular proteasomes (Dianzani et al., 2017) may generate epitopes independently of APC processing.

Involvement of a membrane-bound organelle in the formation of fusion peptides has also been suggested by (Wang et al., 2019). This study utilised the B9-23 epitope which is known to require C-terminal modification for efficient binding and presentation by IA-g7 (Crawford et al., 2011; Stadinski et al., 2010b) and HLA-DQ8 (Nakayama et al., 2015; Yang et al., 2014). The authors advance this theory by identifying potential donor peptides from mouse or human C-peptide for fusion by transpeptidation to insulin B9-23 derived peptides. These fusion peptides create super-agonists, stimulating T cell activation as strongly or if not stronger than mutations to the C-terminal end of the B9-23 peptide. The authors suggest transpeptidation in the lysosome could be responsible for such fusions *in vivo* although proof that transpeptidation of these motifs is operative *in vivo* and produces relevant neo-epitopes is lacking. Speculation remains on the sub-cellular compartment in which transpeptidation occurs as granule purification from whole β-cells is likely to yield crude preparations which may also contain lysosomes.

Alternatively, other mechanisms may exist which also culminate in the generation of non-germline peptide sequences. These include translocation, interstitial deletion and chromosomal

inversion of genes to generate gene fusions documented predominately in cancer cells (Campbell et al., 2008; Mitelman et al., 2007). The identification of novel tumour epitopes generated as a result of gene translocation and point mutations is an exciting research area which may be fruitful in cancer immunotherapy if these epitopes can be selectively targeted (Yarchoan et al., 2017). Indeed, following checkpoint blockade tumour regression is correlated with $CD8^+$ T cells specific for mutation-induced neo-epitopes (Gubin et al., 2014) and these neo-epitopes have been efficacious as vaccines for melanoma (Kreiter et al., 2015). Chimeric RNAs have also been described in cancer cells to generate non-germline sequences by *cis*- and *trans*-splicing of RNA (Jia et al., 2016). Peptide elution from human β-cell lines has identified HLA-A2-restricted neo-antigenic peptides formed by mRNA splice variants of β-cell antigens (Gonzalez-Duque et al., 2018). Thus, there is evolving evidence that such mechanisms occur in autoimmunity and in theory, hybrid peptides could be generated by *trans*-splicing of RNA.

1.8.5 Defective ribosomal products (DRiPs)

During protein homeostasis, proteins can assume the incorrect conformation despite translation from the correct start codon. In this event, these misfolded proteins (defective ribosomal products, DRiPs) traffic to the proteasome where they are degraded and epitopes generated for loading onto class I molecules (Yewdell, 2011). An additional class of DRiPs is generated by translation initiation at an out-of-frame start codon, translating a novel amino acid sequence (Anton and Yewdell, 2014). In tumour and virally infected cells, such pathways play an important role in generating non-conventional class I epitopes for immunosurveillance (Starck and Shastri, 2011).

In the context of T1D, this additional class of DRiPs is involved in the generation of class I and II epitopes (Kracht et al., 2017). Initiation of translation at an alternative start codon within the mRNA of preproinsulin (PPI) defines an alternative open reading frame which bypasses the original PPI stop codon. A neopolypeptide is generated, the amino acid sequence of which is different from PPI and is hence termed a defective ribosomal insulin gene product (INS-DRiP) (Figure 1.4). A 9-mer epitope of this INS-DRiP polypeptide (INS-DRiP$_{1–9}$) is presented by HLA-A2 as well as the highest-T1D-risk HLA-DQ8trans molecules and INS-DRiP$_{1–9}$-specific CD8^{+} T cells shown to kill human β-cells *in vitro*. The expression of the DRiP polypeptide is increased under ER stress with increased β-cell death observed by preconditioning islets with high glucose and pro-inflammatory cytokines (IL-1β and IFN-γ). Whether central tolerance exists against DRiPs, and when responses to the INS-DRiP arise during the course of T1D development, will be vital questions to be addressed.

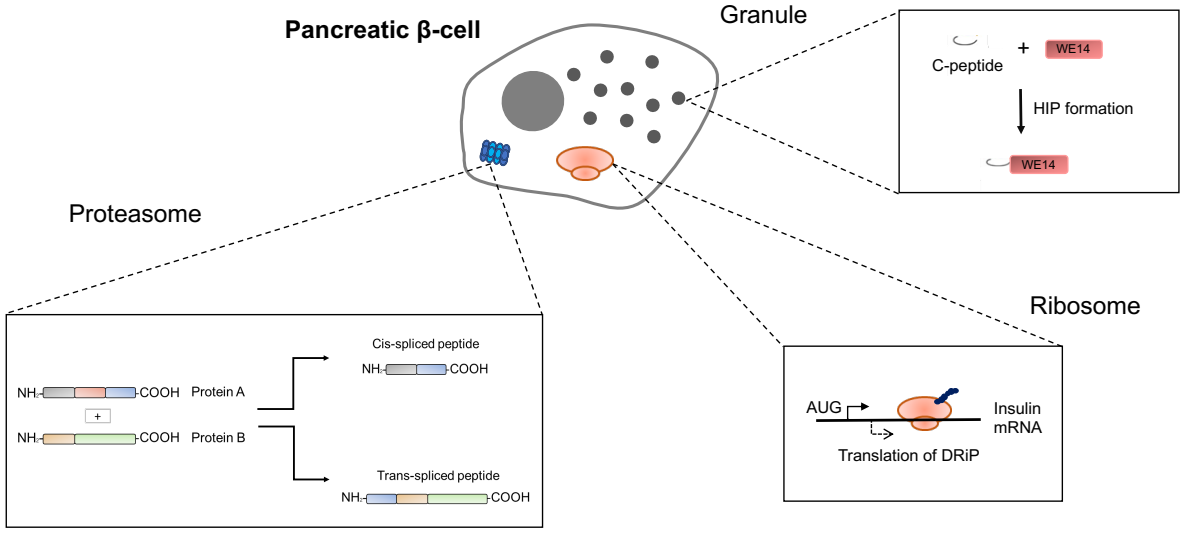


Figure 1.4 Sites of non-conventional epitope generation within the pancreatic β-cell. Epitopes can be generated by errors in translation, peptide splicing by the proteasome and during cleavage of polypeptide cargoes within the granule. DRiP, defective ribosomal product; HIP, hybrid insulin peptide. Taken from (Harbige et al., 2017).

1.9 Implications of non-conventional epitopes on the consensus model

Non-conventional epitope generation perturbs our notion of self and non-self, promoting re-evaluation of the targeting of self-antigens in autoimmune diseases e.g. T1D. A key question is whether hybrid peptides have a physiological role within β-cells which provides a rationale for their generation. The generation of chimeric fusion RNAs, not only in tumour cells but also in non-cancer cells, have been shown to be functional and are required for cellular activities (Babiceanu et al., 2016) although it is yet unclear whether hybrid peptides can be generated via this route. Proteasome-generated spliced peptides have been scrutinised thus far as T cell targets, but they may also have additional functions. For example, spliced peptides may be beneficial for cells to expand the diversity of functional peptides derivable from a single parent protein (Dianzani et al., 2017).

The N- and C-terminal ligation partners involved in proteasome-catalysed peptide splicing appear to be subject to splicing rules. Attempts to identify a HLA-A2-restricted splicing pattern has yielded markedly different results (Berkers et al., 2015; Liepe et al., 2016). (Liepe et al., 2016) eluted spliced peptides from various cell lines and showed substantial differences in the nature of the junctional amino acids, suggesting this is influenced partly by the HLA molecule to which these peptide bind. Equally, (Faridi et al., 2018) report only subtle preferences at the junctional amino acids with no obvious splicing rules. The challenge in identifying splicing rules is likely underpinned by the wide specificity of different proteasome isoforms e.g. constitutive, immunoproteasomes and intermediate proteasomes, which preferentially regulate specific pathways in different cell types (Mishto and Liepe, 2017). Contingent on the expressed isoform, different peptide products can predominant in the peptide repertoire which can determine whether a specific T cell response is primed (Platteel et al., 2017b). Importantly, although not yet established, if presentation of spliced peptides is restricted to specific HLA

alleles, then this has implications for T1D and autoimmune diseases in general. High overlap between self and non-self may facilitate the generation of cross-reactive T cells due to molecular mimicry of pathogen-derived and self-peptides (Cusick et al., 2012).

Cell or tissue-specific restriction of spliced peptides, and whether such epitopes are available for T cell selection in the thymus, will be critical to our understanding of their role in T1D pathogenesis. Low or absent expression of non-conventional epitopes in the thymus may lead to a lack of clonal deletion and thus, escape of non-conventional-reactive T cells into the periphery likely characterised by a more diverse T cell repertoire (Anderson et al., 2000). Within this T cell repertoire, high affinity TCRs to the non-conventional epitope are more likely to exist (Mohan et al., 2010) increasing the opportunity for autoreactive responses. Indeed, MHC II-islet peptide-specific tetramer staining in the NOD mouse has shown that $CD4^{+}CD25^{-}$ (Tconv) T cells display higher reactivity to HIPs than insulin B chain peptides (Holohan et al., 2019). This finding is consistent with the notion that insulin-reactive T cells would undergo thymic deletion while HIPs represent peripherally generated neo-epitopes that escape thymic selection.

Additionally, the modifications present in non-conventional epitopes may also cause a heteroclitic effect whereby modification of key anchor residues enhances peptide binding to HLA molecules and/or TCR docking compared to the conventional epitope (Borbulevych et al., 2005; Cole et al., 2010). Alterations in HLA register usage may also come into play, permitting non-conventional epitopes to be presented as neo-epitopes if in the thymus the epitope was not presented in the same register (Pugliese, 2017), as described for the insulin B chain (Mohan et al., 2011). Regardless of the mechanism, tetramer staining could be useful in pre-diabetic phases to elucidate if these peptides do indeed have the potential to break tolerance.

It is also possible that a reduced pool of thymic-derived Tregs is resident in the periphery due to low or absent expression of non-conventional epitopes in the thymus (Kieback et al., 2016). In support of this, evidence in the NOD mouse suggests that the Treg compartment within the islet is largely comprised of thymically-derived Tregs; islet Tregs isolated from NOD mice do not react with C-peptide:IAPP2 or C-peptide:WE14 HIP tetramers (Holohan et al., 2019). These results suggest that there might not be a substantial proportion of HIP-reactive thymic Tregs resident in the islets overall. It is, however, known that DCs can acquire antigens present in the periphery and traffic them to the thymus to mediate T cell selection (Bonasio et al., 2006; Hadeiba et al., 2012) therefore the thymic Treg repertoire against non-conventional epitopes may be dependent on the stage of disease.

The role of non-conventional epitopes in the context of the disease-protective HLA-DQ6 haplotype is also yet to be determined. HLA-DQ6 may bind these epitopes in a different register than disease-susceptible alleles leading to the induction of Tregs as reported for an epitope of type IV collagen in Goodpasture disease (Ooi et al., 2017). Using HIPs as an example, an increased frequency of HIP-reactive Tregs in individuals with the HLA-DQ6 haplotype would be speculative at this stage, however a recent publication by (Wen et al., 2020) provides some evidence to support this notion. In this study, GAD65- and IGRP-specific effector and $CD25^{+}CD127^{-}FOXP3^{+}$ regulatory $CD4^{+}$ T cells were found to be present in high frequencies in individuals with the protective HLA-DQ6 haplotype than those with susceptible or neutral haplotypes. Alternatively, HLA-DQ6 may bind non-conventional epitopes in a competing HLA-binding register (epitope stealing) than disease-susceptible HLA alleles (van Lummel et al., 2019) directing the T cell response away from the epitope presented by disease-susceptible HLA molecules.

1.10 Aims and significance of the thesis

Efforts to induce immunological tolerance in T1D patients using antigen-specific therapies e.g. peptide immunotherapy (Alhadj Ali et al., 2017) have focused on the use of conventional epitopes. The studies discussed above highlight the relevance of non-conventional epitopes in T1D and utilising HIPs for tolerance induction has already shown promise in the NOD mouse (Jamison et al., 2019). In order to understand whether non-conventional epitope generation can be targeted at a therapeutic level, this requires a more comprehensive understanding of peptide generation. For example, the pathways through which hybrid peptides are loaded onto HLA class II molecules are not clear. Equally, hybrid peptides have thus far been identified by proteolytic digestion of β-cell granule extracts prior to mass spectrometric analysis. It is therefore unclear whether the granule extract contains these hybrid peptide species or whether they derive from a larger species e.g. a hybrid polypeptide that requires immunological processing. This is an important question since a better understanding of the antigen processing pathway and its cellular location would potentially foster novel therapeutic approaches to countering these events.

The hypothesis to be tested in this thesis is that epitopes from hybrid polypeptides undergo immunological processing by conventional antigen-presenting cells and presentation by high-risk HLA molecules. The aims of the study are given below with the experimental objectives:

1. **Identify novel naturally processed and presented epitopes (NPPEs) derived from hybrid polypeptides generated by the HLA class II processing pathway**
    - Optimisation of an antigen delivery system (ADS)
    - APC pulsing and pHLA elution to identify novel hybrid epitopes

2. **Investigate immunogenicity of hybrid polypeptides**

    - Peripheral blood T cell reactivity against hybrid polypeptides

CHAPTER 2: MATERIALS AND METHODS

2.1 Antigens and peptides

All peptides were obtained lyophilised and re-constituted in dimethyl sulfoxide (DMSO). Peptides utilised in antigen delivery experiments and the synthetic pokeweed peptide used to verify the spectra of the natural peptide, were purchased from Thermo Fisher Scientific (Rockford, IL, USA) or Almac Sciences (Edinburgh, Scotland). Peptides were synthesised by Fmoc (9-fluorenylmethoxycarbonyl) solid-phase chemistry to more than 95% purity using mass spectrometry as well as reversed-phase high-performance liquid chromatography (RP-HPLC) to confirm identity and for purification of crude peptides. The amino acid sequence of each peptide is shown in Table 2.1. The control peptide contains the HLA-DR4-restricted GAD65(274-286) epitope embedded into the intersection of the C-peptide:WE14 sequence. 41-mer peptides were synthesised with an added cysteine residue at the C-terminus and a biotin moiety at the N-terminus.

Table 2.1 Amino acid sequence of synthetic pokeweed peptide and peptides used for antigen delivery.

Peptide	MW (g/mol)	Amino acid sequence
GAD65(274-286)	1504	IAFTSEHSHFSLK
C-peptide:WE14	4467	Biotin–*EAEDLQVGQVELGGGPGAGSLQPLAL*WSKMDQLAKELTAEC
Control	4584	Biotin–*GGPGAGSLQPLAL*IAFTSEHSHFSLKWSKMDQLAKELTAEC
PWM lectin-B (47-58)	1289	REASGKVCPDDL

†Highlighted sequence represents GAD65(274-286) epitope. Italicised and underlined sequences represent C-peptide and WE14, respectively. PWM, pokeweed mitogen; MW, molecular weight.

Peptides used for FluoroSpot assays were synthesised by GL Biochem (Shanghai, China) to more than 95% purity using Fmoc-Glu(OtBu)-Wang resin (Table 2.2). Infanrix-hexa, a hexavaccine consisting of diphtheria, tetanus, pertussis, inactivated poliomyelitis, *Haemophilus influenzae* type b and hepatitis B was obtained from GlaxoSmithKline and used at 2 µL/mL per well as a positive control. Lyophilised *Candida albicans* antigen was purchased from Greer Laboratories (strain #10231, cat no: XPLM73X1A2), re-constituted to 1 mg/mL in phosphate buffered saline (PBS) and used at a final concentration of 2 µg/mL as a positive control for IL-17A responses.

Table 2.2 List of peptides used for FluoroSpot assays.

Peptide	Amino acid sequence
hEL:ChgA-WE14	SLQPLALWSKMDQL
C-peptide:WE14	EAEDLQVGQVELGGGPGAGSLQPLALWSKMDQLAKELTAE
WE14:C-peptide	WSKMDQLAKELTAEEAEDLQVGQVELGGGPGAGSLQPLAL
hEL:IAPP2	SLQPLALNAVEVLK
C-peptide:IAPP2	EAEDLQVGQVELGGGPGAGSLQPLALNAVEVLKREPLNYLPL
IAPP2:C-peptide	NAVEVLKREPLNYLPLEAEDLQVGQVELGGGPGAGSLQPLAL

†Highlighted sequences represent C-peptide components of HIPs. Blue and red represent WE14- and IAPP2-containing sequences, respectively. HIP, hybrid insulin peptide.

2.2 Cell lines

Table 2.3 lists the Epstein–Barr virus (EBV)-transformed B-lymphoblastoid cell lines (BLCLs) used in this study obtained from Public Health England (PHE)/European Collection of Authenticated Cell Cultures (ECACC). B cell lines were maintained in R10 medium (RPMI-1640 GlutaMAX [Thermo Fisher Scientific], 100 U/mL penicillin, 100 µg/mL streptomycin [penicillin-streptomycin 100X stock solution, Thermo Fisher Scientific] and 10% [v/v] heat-inactivated FCS) at a density of 0.3–0.6 × 10^6 cells/mL. To heat-inactivate, FCS was heated at 56°C for 30 minutes in a water bath with occasional shaking.

Table 2.3 EBV-transformed B-lymphoblastoid cell lines.

Cell line	DNA typing profile		
	DRB1*	***DQA1****	***DQB1****
WT51	*04:01*	*03:01*	*03:02*
PF04015	*03:01*	*05:01*	*02:01*

The murine T cell hybridoma T33 was donated by Dr Linda Wicker (University of Cambridge, Cambridge, UK). The hybridoma was generated in a HLA-DR4 transgenic mouse and recognises the GAD65(274-286) epitope presented by HLA-DR4 (Wicker et al., 1996). T33 hybridoma cells were seeded in D10 medium (Dulbecco's Modified Eagle Medium [Thermo Fisher Scientific], 100 U/mL penicillin, 100 µg/mL streptomycin and 10% [v/v] heat-inactivated FCS) at a density of 0.1×10^6 cells/mL and split 1:10 every 2-3 days.

2.3 Chapter 3

2.3.1 DEC205 surface expression

For all experiments, cells were used when split the day prior to assure good quality. WT51 cells were harvested and washed with 4 mL of FACS buffer (FB) comprised of 1X PBS and 3% (v/v) FCS and centrifuged (300*g*, 5 minutes, 20°C). After washing the tube for staining, the supernatant was decanted, cells were re-suspended in 100 µL of FB and 5 µL of Human TruStain FcX Fc receptor blocking solution (BioLegend, UK) added for 10 minutes on ice. WT51 cells were subsequently incubated with 2 µL of anti-human DEC205 antibody (BD Biosciences, PE-conjugated IgG2b, MG38 clone) or corresponding isotype control on ice for 30 minutes in the dark. Cells were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 20°C) and the supernatant discarded. A 100 µL residual volume remained and 1 µL of 7-AAD Viability Staining Solution (eBioscience) was added for 5 minutes at room temperature in the

dark before analysis on a BD FACSCanto (BD Biosciences, San Jose, CA, USA). Unless otherwise stated, flow cytometry was performed by gating on live single cells (Figure 2.1) and FACS data analysed using FlowJo 10.1 software (FlowJo LLC).

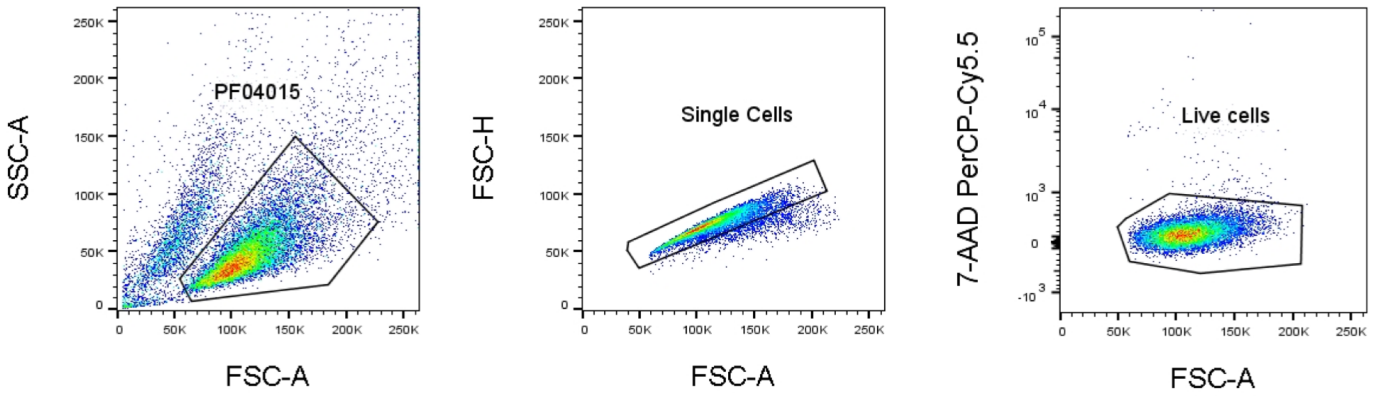


Figure 2.1 Representative gating strategy of FACS analysis. Gating was performed by first gating for the cell population based on forward and side scatter (far left). Doublets were then excluded and 7-AAD negative live cells gated (far right).

2.3.2 Generation of pokeweed mitogen and anti-DEC205 conjugates

Purified mouse anti-human DEC205 antibody was purchased (Bio-Rad, non-conjugated IgG2b, MG38 clone, cat no: MCA2258EL) at 1 mg/mL in PBS. Lectin from *Phytolacca americana* (pokeweed mitogen) was obtained from Sigma-Aldrich (cat no: L9379) as a lyophilised powder and re-constituted in PBS to 1 mg/mL. The extinction co-efficient of pokeweed mitogen ($E^{1\%}$ = 18.5) was used to accurately determine the protein concentration by a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific).

Prior to conjugation, both proteins were prepared in Conjugation Buffer (1X PBS + 2 mM EDTA [0.5 M stock, Sigma-Aldrich]). Pokeweed mitogen and anti-DEC205 antibody were treated with a molar excess of sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1 carboxylate (10 mg/mL stock solution in double-distilled water; sulfo-SMCC from Thermo Fisher Scientific) crosslinker. Depending on the protein concentration, the molar excess of

crosslinker ranged from twenty-fold (protein concentration of 1-4 mg/mL) to fifty-fold (protein concentration of <1 mg/mL). After 30 minutes of incubation at room temperature with moderate agitation on an orbital shaker, excess crosslinker (MW: 436.37 g/mol) was removed using 7K MWCO Zeba Desalting Columns (Thermo Fisher Scientific). The concentration of sulfo-SMCC activated carrier protein was determined after desalting using a NanoDrop spectrophotometer by measuring absorbance at 280 nm. The instrument uses the Beer-Lambert law ($A = \varepsilon\ c\ l$; where A is absorbance, ε is molar extinction co-efficient, c is concentration and l is light path length) with the appropriate extinction co-efficients (IgG: 13.7 and pokeweed mitogen: 18.5) to determine protein concentration. Subsequently, peptide containing a C-terminal cysteine and a N-terminal biotin group was added to sulfo-SMCC activated proteins. Peptides were added in Conjugation Buffer at a peptide/protein molar ratio of 5:1 (activated anti-DEC205 antibody; MW: 150 kDa) or 2:1 (activated pokeweed mitogen; MW: 32 kDa). Following 30 minutes of incubation at room temperature with moderate agitation on an orbital shaker, the generated conjugates were stored at -20°C until use. After thawing, conjugates were stored at 4°C.

2.3.3 Pulsing of EBV-transformed B cells with crosslinked GAD65(274-286)-containing control peptide

B cell lines were seeded at 10×10^6 cells/mL in R10 in sterile 1.5 mL Eppendorf tubes and pulsed with conjugates at a concentration of 2.18 or 0.218 μM for 30 minutes on ice. As a control, at the same time cells were pulsed with activated pokeweed mitogen or anti-DEC205 antibody or peptide (biotin-labelled) alone comparable with the amount present in the conjugate. Where indicated, control EBV-transformed B cells were also pulsed with the equivalent concentration of DMSO peptide diluent or non-conjugated activated pokeweed mitogen and GAD65(274-286)-containing 41-mer. After 30 minutes, R10 (37ºC) was added to

adjust the cell concentration to 0.5×10^6 cells/mL (final ADS concentration of either 0.1 or 0.01 μM) and cells incubated at 37ºC, 5% CO_2 for 16 hours in a 12-well plate (Corning) to permit uptake. In a few experiments, non-pulsed cells were seeded in R10 at 0.5×10^6 cells/mL and three concentrations (2.18, 0.218 and 0.0218 μM) of control peptide (41-mer containing GAD65[274-286] epitope) added directly to the culture for 16 hours at 37ºC, 5% CO_2.

EBV-transformed B cells were harvested, washed with 8 mL of R10 and centrifuged (300*g*, 5 minutes, 20°C). The supernatant was discarded and cells were counted (using Trypan blue [Sigma-Aldrich]) before co-culture with T33 hybridoma cells at a B cell-T cell ratio of 1:1. In each case, 100 μL of cells were added per well at a concentration of 1×10^6 cells/mL in a 96-well plate (200 μL final volume per well). As a control, WT51 cells (DR4/DQ8) were pulsed at 1×10^6 cells/mL (100 000 cells/100 μL) for 30 minutes at 37ºC with the GAD65(274-286) epitope before subsequent co-culture with the T33 hybridoma (1:1 B cell-T cell ratio as above). GAD65(274-286) was used at equimolar (2.18 μM) to that of the 41-mer peptide in the ADS. T33 cells cultured with untreated EBV B cells and both cell lines cultured alone were used at negative controls. Following 6 hours of co-culture at 37°C, 5% CO_2, supernatants were harvested for an IL-2 ELISA. IL-2 production was evaluated using the Mouse IL-2 ELISA Ready-SET-Go Kit (eBioscence, San Diego, CA, USA) according to the manufacturer's instructions.

2.3.4 Detection of conjugates at the cell surface

EBV-transformed B cells were pulsed in multiples of 10×10^6 cells/mL as in (2.3.3) with 2.18 μM of pokeweed-peptide conjugate, equivalent concentrations of activated pokeweed or peptide (biotin-labelled) alone and incubated at 37°C, 5% CO_2 for 0–6 hours before being placed on ice to stop further internalisation. Samples were washed with 4 mL of FB, centrifuged

(300*g*, 5 minutes, 4°C) and the supernatant taken off. Cells were re-suspended in 100 μL of FB and 5 μL of Human TruStain FcX Fc receptor blocking solution added for 10 minutes on ice prior to the addition of 1 μL of anti-biotin antibody (BioLegend, PE-conjugated IgG2a, 1D4-C5 clone) or corresponding isotype control. Following 30 minutes of incubation on ice in the dark, cells were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 4°C) and the supernatant decanted. 1 μL of 7-AAD Viability Staining Solution was added in 100 μL suspension volume for 5 minutes at room temperature in the dark before acquisition on a flow cytometer. Cells pulsed with 2.18 μM of anti-DEC205-peptide conjugate were incubated at 37°C, 5% CO_2 over a shorter time period (0–2 hours) before biotin surface staining. In this case, staining was performed as above except 1 μL of APC-conjugated streptavidin (1 mg/mL; Invitrogen) in 100 μL was used to detect biotinylated peptide at the cell surface after pulsing.

In some experiments it was necessary to confirm binding of the anti-DEC205-peptide conjugate to the cell surface. WT51 cells were pulsed at 10×10^6 cells/mL in R10 for 30 minutes on ice with either non-activated anti-DEC205, sulfo-SMCC activated anti-DEC205 or anti-DEC205-peptide conjugate ensuring incubation with the same amount of antibody in each condition. Control cells were also pulsed with a mouse IgG2b negative control antibody (Bio-Rad, cat no: MCA691). After washing with 4 mL of FB and adding Fc receptor block solution, 1 μL of anti-mouse IgG2b antibody (BioLegend, PE-conjugated IgG1, RMG2b-1 clone) was added in 100 μL staining volume for 30 minutes on ice in the dark. Samples were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 4°C) and the supernatant discarded. 1 μL of 7-AAD Viability Staining Solution was added in 100 μL suspension volume for 5 minutes at room temperature in the dark before acquisition on a BD FACSCanto flow cytometer.

2.3.5 Detection of biotinylated peptide within anti-DEC205 conjugate

Anti-mouse Ig κ CompBeads (BD Biosciences, San Jose, CA) were washed once with 4 mL of FB, spun at 300*g* for 5 minutes (20°C) and the supernatant discarded before addition of 2.18 μM of anti-DEC205-peptide conjugate for 20 minutes at room temperature. As a control, CompBeads were also incubated with activated anti-DEC205 antibody and peptide alone equal to that present in the conjugate. After 20 minutes beads were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 20°C) and the supernatant discarded before adding 1 μL of APC-conjugated streptavidin for 30 minutes on ice in the dark. Samples were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 4°C) and the supernatant decanted prior to acquisition on a flow cytometer by gating on the bead population.

2.3.6 Imaging flow cytometry

WT51 cells were pulsed at 10×10^6 cells/mL with either 10 μg/mL of fluorophore-conjugated anti-human DEC205 antibody (BD Biosciences, PE-conjugated IgG2b, MG38 clone) or 5 μg/mL of APC-conjugated pokeweed mitogen for 30 minutes on ice. Pokeweed mitogen was pre-conjugated with APC using an APC conjugation kit (Abcam ab201807), following the provided protocol. Cells were also pulsed with either an anti-human HLA-DR antibody (BioLegend, APC-conjugated IgG2a, L243 clone) or an anti-human HLA-DQ antibody (BioLegend, PE-conjugated IgG1, HLADQ1 clone) as a positive control for surface staining. After pulsing, cells were seeded at 0.5×10^6 cells/mL in R10 at 37°C, 5% CO_2 in a 12-well plate. At specific time points as indicated in the figure legend, cells were harvested into sterile 1.5 mL Eppendorf tubes and placed on ice. Samples were washed three times with 1 mL of FB (4°C), centrifuged (300*g*, 5 minutes, 4°C) and the supernatant removed by aspiration. Cells were re-suspended in 50 μL of FB (4°C) and analysed on the ImageStream X MKII (Amnis-Luminex) imaging flow cytometer at 40X magnification. A minimum of 20 000 pre-gated cells

were acquired per sample. Channel 1 (PE-conjugated anti-DEC205 monoclonal antibody) or 9 (APC-conjugated pokeweed mitogen) detects brightfield. Channel 11 and 3 was used for APC and PE, respectively. Channel 6 was used for side scatter. All other channels were turned off during acquisition. Laser powers were set to the maximum.

For data analysis, Amnis® IDEAS version 6.2 (Luminex Corporation) was used. Cells were analysed using the gating strategy shown in Figure 2.2: single cells (brightfield area versus aspect ratio), cells in focus (histogram of gradient RMS) and cells positive for either APC or in some cases PE. For the measurement of internalisation, an object mask based on the brightfield cell image (channel 1/9) was first created and then eroded by 4-5 pixels to exclude the cell membrane. The internalisation score was then calculated on this eroded object mask as well as the fluorescent image of the relevant channel in the Internalisation feature provided in the IDEAS software. The internalisation score is defined as the ratio of the intensity inside the cell to the intensity of the entire cell. Samples pulsed for 0 hours were used as imaging controls.

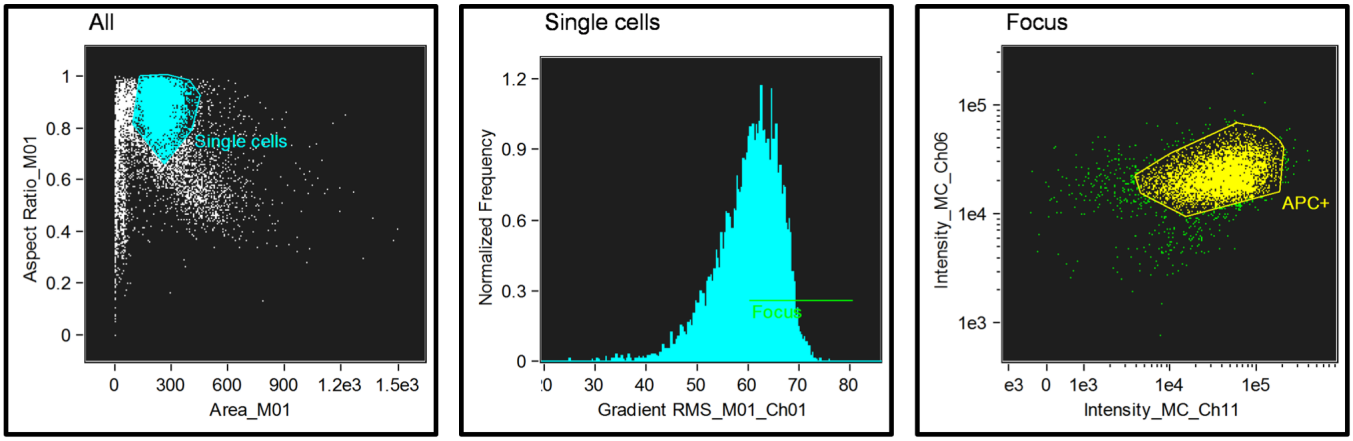


Figure 2.2 Gating strategy for ImageStream analysis. Single cells were first identified based on a dot plot of brightfield area versus aspect ratio (far left). Using this population, cells are then gated based on focus using gradient RMS (middle). To gate PWM-APC-positive cells, a dot plot of the side scatter intensity (Intensity_Ch06) versus the fluorescence intensity of APC (Intensity_Ch11) was used (far right). APC+ gate was set using an unstained control. This population was used for all downstream analyses.

2.3.7 Fixation of EBV-transformed B cells pulsed with crosslinked GAD65(274-286)-containing control peptide

EBV-transformed B cells were pulsed with 2.18 μM of pokeweed-peptide conjugate or activated pokeweed mitogen alone as described previously (2.3.3) and incubated at 37°C, 5% CO_2 for specific time periods (indicated in figure legend). After being removed from culture cells were washed once with 1X PBS (no serum), centrifuged (200*g*, 3 minutes, 20°C) and the supernatant removed before gentle fixation with 2 mL of 0.1% (v/v) paraformaldehyde in PBS (16% stock solution, Electron Microscopy Sciences) for 30 minutes at room temperature in the dark. Samples were then washed twice with 1X PBS-5% (v/v) FCS, centrifuged (200*g*, 3 minutes, 20°C) and the supernatant decanted. Cells were counted and cultured at a 1:1 ratio (100 000 cells each) in a 96-well plate with the T33 hybridoma at 37°C, 5% CO_2. T33 cells were also cultured alone or with PMA (50 ng/mL) and ionomycin (1 μg/mL) as negative and positive controls, respectively. Supernatants were harvested after 16 hours of incubation for determination of IL-2 as mentioned previously (2.3.3).

2.4 Chapter 4

2.4.1 Pulsing of EBV-transformed B cells with pokeweed mitogen crosslinked to C-peptide:WE14

For each pulsing experiment, 100–150 × 10^6 B cells were pulsed at the same time, with multiple rounds of pulsing being performed. EBV-transformed B cells were centrifuged at 500*g* for 8 minutes at room temperature. All subsequent centrifugation steps were performed at 4°C. The cell pellet was washed first with 50 mL of room temperature 1X PBS and then with 1X PBS (4°C) by centrifugation at 400*g* for 5 minutes. Cells were re-suspended in 1X PBS (4°C) to 20 × 10^6 cells/mL and placed on ice. Crosslinked pokeweed-hybrid polypeptide was diluted to 4.36 μM in 1X PBS (4°C) and mixed in a 1:1 ratio with the cell suspension giving an ADS final

concentration of 2.18 μM. Following 30 minutes of incubation on ice, R10 medium (37°C) was added to adjust cells to 0.5×10^6 cells/mL (final ADS concentration 0.1 μM) and seeded in multiple T175 flasks (Corning) for 16 hours laid flat in the incubator at 37°C, 5% CO_2. Pulsed cells were harvested, pooled and washed twice with 50 mL of 1X PBS (4°C) by centrifugation (400*g*, 5 minutes, 4°C). A final wash step was performed (600*g*, 5 minutes, 4°C) to give a dense cell pellet. The supernatant was aspirated and the cell pellet stored at -80°C for immunoaffinity purification.

2.4.2 Anti-biotin and HLA class II surface staining of EBV-transformed B cells pulsed with crosslinked C-peptide:WE14

After pulsing with the pokeweed-C-peptide:WE14 conjugate, EBV-transformed B cells were stained with an anti-biotin antibody following the protocol described previously (2.3.4). To detect surface HLA-DR and -DQ expression, Fc block was performed and cells stained for 30 minutes on ice with 10 μg/mL of anti-DQ antibody (Bio-Rad, non-conjugated IgG2a, SPVL3 clone, cat no: MCA279G) or negative control mouse IgG2a antibody (Dako Cytomation, cat no: X0943). After washing twice with 4 mL of FB, cells were stained for 30 minutes on ice in the dark with 1 μL of anti-mouse IgG2a secondary antibody (BioLegend, PE-conjugated IgG, RMG2a-62 clone) and 1 μL of anti-human HLA-DR antibody (BioLegend, APC-conjugated IgG2a, L243 clone). Samples were also stained with the corresponding HLA-DR isotype control antibody. Afterward, cells were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 20°C) and the supernatant discarded. 1 μL of 7-AAD Viability Staining Solution was added in 100 μL for 5 minutes at room temperature in the dark before acquisition on a BD FACSCanto flow cytometer.

2.4.3 Purification of anti-DR (L243) and anti-DQ (SPVL3) antibodies

Murine ascites fluid containing the L243 monoclonal antibody was generated previously by Harlan Sera-Lab (Loughborough, UK). SPVL3 ascites was produced by ProMab Biotechnologies (Richmond, CA, USA) by firstly priming BALB/c mice with 0.5 mL pristine for 1-2 weeks before injecting 1-2 × 10^6 viable hybridoma cells per mouse. The ascites fluid was harvested after 4 weeks and centrifuged to remove particles. L243 and SPVL3 antibodies were purified from ascitic fluid using the Melon Gel Monoclonal IgG Purification Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, the ascites fluid was treated with Ascites Conditioning Reagent (Thermo Fisher Scientific, cat no: 45219) before dialysing into 1X Melon Gel Purification Buffer. The dialysed sample is added to a column containing a proprietary resin which binds non-IgG proteins and pure IgG is recovered in the flow through fraction. Purified antibodies were buffer-exchanged into PBS for downstream applications using 7K MWCO Zeba Delating Columns. The concentration and purity of the recovered antibodies were analysed by a NanoDrop spectrophotometer and SDS-PAGE, respectively.

2.4.4 Confirmation of HLA binding capacity of in-house purified antibodies

WT51 cells were washed with 4 mL of FB, centrifuged (300*g*, 5 minutes, 20°C) and the supernatant removed before re-suspending in 100 µL of FB and incubating with 5 µL of Human TruStain FcX Fc receptor blocking solution for 10 minutes on ice. Cells were stained with a titration (0.01 – 10 µg/mL) of either in-house purified antibodies, commercially available L243 (LEAF Purified anti-human HLA-DR Antibody, BioLegend)/SPVL3 (Anti-HLA-DQ antibody [SPV-L3] ab23632, Abcam) or 10 µg/mL of negative control mouse IgG2a antibody for 30 minutes on ice. After washing twice with 4 mL of FB, samples were stained with 1 µL of PE-conjugated anti-mouse IgG2a secondary antibody for a further 30 minutes on ice in the dark.

Cells were washed twice with 4 mL of FB, centrifuged (300*g*, 5 minutes, 20°C) and the supernatant decanted before acquisition on a BD FACSCanto flow cytometer.

2.4.5 Crosslinking of antibodies to Protein A Sepharose

One millilitre of Protein A Sepharose beads (Fast Flow; GE Healthcare) was added to a Poly-Prep Chromatography column (BioRad), washed with 10 resin volumes (RV) of a solution containing 0.1M boric acid and 0.1M KCl (borate buffer; pH 8.0) and subsequently incubated with 5 (SPVL3) to 10 mg (L243) of purified anti-HLA-class II antibody with moderate rotation for 1 hour at room temperature. The beads were then washed with 10 RV of 0.2 M triethanolamine (pH 8.2; Sigma-Aldrich) and the bound antibody was crosslinked by using 40 mM dimethyl pimelimidate dihydrochloride (DMP; Sigma-Aldrich) (pH 8.3) for 1 hour at room temperature. The crosslinking reaction was terminated by the addition of 10 RV ice-cold 0.2 Tris buffer (pH 8.0). Any unbound antibody was removed by using 0.1M citrate buffer (pH 3.0) and the column stored in borate buffer supplemented with 0.02% (w/v) sodium azide (Sigma-Aldrich) at 4°C until use.

2.4.6 Isolation of peptide-HLA complexes and peptide purification

Cell pellets were thawed on ice and dissociated using 2 mL (per 100×10^6 cells) of cell lysis buffer (150 mM NaCl, 20 mM Tris [pH 8.0], 1 mM EDTA, 6 mM Sodium deoxycholate detergent [Sigma-Aldrich], 34 mM Octyl β-D-glucopyranoside [Sigma-Aldrich] and protease inhibitors [Complete Protease Inhibitor Cocktail tablets, Roche Life Science, Indianapolis, IN]). Cell membranes were further disrupted by sonication (QSonica, Newtown, CT) using five (100 ms/s) pulses at 20% amplitude until all the visible precipitates were solubilised. Lysates were pre-cleared using microfuge centrifugation (20 000*g*, 30 minutes, 4°C). Soluble cell lysates were loaded by gravity flow onto a series of pre-equilibrated immunoaffinity columns

set up in tandem at 4°C. To remove non-specific binding proteins, the lysate was first passed over a column containing Protein A Sepharose beads without crosslinked antibody. The flow through from this pre-column was set up to drip directly onto affinity columns with crosslinked L243 and SPVL3 antibodies. Lysate loading onto the affinity columns was repeated three times before separating the columns for washing steps of 20 RV with low salt buffer (150 mM NaCl, 20 mM Tris, pH 8.0), high salt buffer (450 mM NaCl, 20 mM Tris, pH 8.0), low salt buffer and no salt buffer (10 mM Tris, pH 8.0). The peptides were subsequently eluted with 5 RV of 10 % (v/v) acetic acid (Acetic Acid, Glacial Optima LC/MS Grade [Fisher Chemical]).

After elution, HLA molecules and HLA-binding peptides were separated using a Sep Pak C18 cartridge (Waters, Milford, MA, USA). The cartridge was wetted twice with 1 mL 50% (v/v) acetonitrile (ACN):double-distilled water (ddH_2O) and rinsed twice with 1 mL 0.1% (v/v) trifluoroacetic acid (TFA):ddH_2O. After equilibrating the cartridge twice with 1 mL 2% (v/v) ACN:0.1% TFA:ddH_2O, the sample was loaded onto the cartridge and the flow through collected. The cartridge was washed twice with 1 mL 2% (v/v) ACN:0.1% TFA:ddH_2O and peptides eluted with 400 μL 30% (v/v) ACN:0.1% TFA:ddH_2O into a Protein LoBind Eppendorf tube (Sigma-Aldrich). HLA molecules were eluted using 400 μL 80% (v/v) ACN:0.1% TFA:ddH_2O. Eluates were vacuum concentrated to complete dryness before being delivered to the Proteomics Core Facility at Denmark Hill (King's College London).

2.4.7 Liquid chromatography-tandem mass spectrometry (LC-MS/MS) method and database search

Samples were re-constituted in 2% (v/v) ACN-0.1% formic acid (FA) loading buffer and approximately 50% used for LC-MS/MS analysis. Chromatographic separation was performed using an Ultimate 3000 NanoLC system (Thermo Fisher Scientific). Peptides were resolved by

reversed phase chromatography on a 75 μm × 50 cm C18 column using a four-step linear gradient of water in 0.1% (v/v) formic acid (A) and 80% (v/v) acetonitrile in 0.1% formic acid (B). The gradient was delivered to elute the peptides at a flow rate of 250 nL/minutes over 90 minutes. The eluate was ionised by electrospray ionisation using an Orbitrap-Fusion-Lumos (Thermo Fisher Scientific) operating under Xcalibur v4.1. The instrument was programmed to acquire using a "Universal" method by defining a 3s cycle time between a full MS scan and MS/MS fragmentation. The MS/MS analyses were conducted using collision energy profiles that were chosen based on the mass-to-charge ratio (m/z) and the charge state of the peptide.

Raw mass spectrometry data were processed into peak list files using both Proteome Discoverer (ThermoScientific; v2.2) (PD 2.2) and Peakstudio 7.5. Processed data was then searched using both Mascot search algorithm (www.matrixscience.com) embedded in PD 2.2 and Peaks DB search, against the current version of the reviewed Swissprot Homo sapiens database downloaded from Uniprot (http://www.uniprot.org/uniprot/) combined with in-house created sequences of the C-peptide:WE14 polypeptide and synthetic pokeweed peptide (Table 2.1).

2.4.8 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot

Samples were boiled at 95°C in Laemmli sample buffer (Bio-Rad) containing 355 mM β-mercaptoethanol (Sigma-Aldrich) for 10 minutes and loaded alongside a Color Prestained Protein Standard (Broad Range 10-250 kDa, New England Biolabs) on denaturing polyacrylamide gels (either 10% or 4-20% w/v) purchased from Bio-Rad. Gels were run in a mini-gel apparatus (Bio-Rad) in Tris/Glycine/SDS running buffer (Bio-Rad) at 100V until the dye front reached the end of the gel. After electrophoresis, gels were either used for immunoblotting or stained with Coomassie blue (Coomassie Brilliant Blue R-250 Staining

Solution, Bio-Rad) overnight with moderate agitation. Gels were de-stained in several changes of 10% (v/v) acetic acid (Sigma-Aldrich) until the background was clear.

For Western blotting, gels were transferred onto a nitrocellulose membrane (GE Healthcare) and blocked with 5% (w/v) nonfat dry milk in Tris Buffered Saline (TBS) with 0.1% (v/v) Tween-20 (TBST) for 1 hour at room temperature. The membrane was incubated with either an anti-HLA-DR primary antibody (1:200 dilution; TAL 1B5, Santa Cruz Biotechnology) or an anti-HLA-DQA1 antibody (1:10 000 dilution; EPR7300, ab128959, Abcam) at 4°C overnight. After washing with TBST, a goat anti-mouse or anti-rabbit secondary antibody conjugated to a DyLight 680 fluorescent dye (1:10 000 dilution; Cell Signalling) was added at room temperature for 1 hour. Membranes were washed with TBS and the antibody signals imaged using the Li-Cor Odyssey CLx imaging system.

2.5 Chapter 5

2.5.1 Subjects

Cryopreserved peripheral blood mononuclear cell (PBMC) samples from twenty T1D patients (age range 18-38 yrs, median: 26 yrs; disease duration range 1-11 months, median: 5 months) were obtained from the T1D Biobank (Dr Timothy Tree, King's College London). The Biobank contains a collection of PBMCs from patients considered to have T1D by the local diabetes team (based upon clinical presentation and/or the presence of islet autoantibodies) with an intention to start or has already started on insulin therapy. Diagnosis made within the last 12 months.

Patients were recruited onto the "T cell studies in Type 1 diabetes" study at Guy's and St Thomas' Hospital and Bristol Royal Infirmary. This study was carried out with the approval of

the UK National Research Ethics Service (REC reference 08/H0805/14) and informed consent was obtained from all participants. PBMCs were isolated from fresh heparinised blood on density gradients before being frozen in heat-inactivated 90% foetal calf serum (FCS; Gibco)/10% (v/v) DMSO (Sigma-Aldrich) or in some cases CryoStor (Sigma-Aldrich) and samples stored in liquid nitrogen. Basic demographic information of each patient is shown in Table 2.4. *HLA-DRB1* and *-DQB1* genotyping was performed by the Clinical Transplantation Laboratory (Guy's Hospital, London, UK).

Table 2.4 Patient demographics.

Patient	Sex	Age (y)	T1D (m)	*DRB1**		*DQB1**	
				Allele 1	Allele 2	Allele 1	Allele 2
G695	M	20	9	*03:01*	*13:02*	*02:01*	*06:04*
G697	F	24	6	*03:01*	*04:01*	*02:01*	*03:02*
G705	M	27	9	*04:01*	*13:01*	*03:02*	*06:03*
G712	F	18	5	*03:01*	*04:01*	*02:01*	*03:02*
G713	F	23	8	*04:01*	*04:01*	*03:02*	*03:02*
G716	F	18	1	*04:01*	*13:01*	*03:02*	*06:03*
G729	M	24	3	*04:01*	*13:01*	*03:01*	*06:03*
G732	F	30	5	*03:01*	*03:01*	*02:01*	*02:01*
G733	M	32	6	*07:01*	*13:02*	*02:01*	*06:04*
G772	M	26	3	*03:01*	*04:01*	*02:01*	*03:02*
G804	M	26	4	*04:01*	*13:02*	*03:02*	*06:04*
G817	F	20	<1	*04:01*	*13:01*	*03:02*	*05:01*
G819	M	27	11	*07:01*	*15:01*	*02:01*	*06:02*
G832	M	26	2	*03:01*	*07:01*	*02:01*	*02:01*
NDB115	M	26	1	*04:01*	*04:01*	*03:01*	*03:02*
NDB116	F	18	3	*01:01*	*04:01*	*03:01*	*05:01*
NDB118	M	24	1	*01:01*	*03:01*	*02:01*	*05:01*
NDB119	M	38	7	*04:01*	*13:02*	*03:02*	*06:04*
NDB121	M	30	8	*01:01*	*04:01*	*03:01*	*05:01*
NDB129	F	26	9	*03:01*	*08:01*	*02:01*	*04:01*

[†]A high-resolution typing result is given for *DRB1*04:01*. Other results reported to four digits in this table represent the first allele in a string of alleles, as only low to medium resolution typing was performed. F, female; M, male; G, patients recruited at Guy's Hospital; NDB, patients recruited at Bristol Royal Infirmary. Patient age was rounded down.

2.5.2 FluoroSpot assay

FluoroSpot assays were carried out using an IFN-γ, IL-22, IL-17A FluoroSpot kit (FSP-011803-10, Mabtech, Sweden). Cryopreserved PBMCs from T1D patients were thawed and washed twice with thawing media (RPMI-1640 GlutaMAX, 100 U/mL penicillin, 100 µg/mL streptomycin, 2.5 µg/mL amphotericin B [Fungizone, Thermo Fisher Scientific] and 10% [v/v] heat-inactivated FCS) before centrifugation (first wash: 200*g*, 10 minutes, 20°C; second wash: 400*g*, 5 minutes, 20°C). Cells were seeded in a 48-well plate at a density of 2×10^6 cells/mL in 0.5 mL complete media (RPMI-1640 GlutaMAX, 100 U/mL penicillin, 100 µg/mL streptomycin and 2.5 µg/mL amphotericin B) supplemented with 10% (v/v) heat-inactivated human AB serum (Sigma-Aldrich) supplemented with peptide at a final concentration of 6.64 µM. Control wells contained PBMCs stimulated with an equivalent concentration of DMSO peptide diluent or Infanrix-hexa and *Candida albicans* antigen. After 48 hours, non-adherent cells from each well (1×10^6/well) were harvested using complete media/2% (v/v) human AB serum, centrifuged (400*g*, 10 minutes, 20°C) and the supernatant decanted. Cells were re-suspended in complete media/10% (v/v) human AB serum at a concentration of 3.3×10^6/mL and 100 µL of cells aliquoted into triplicate wells of a pre-coated FluoroSpot plate (mAbs 1-D1K, 9D7 and MT44.6) pre-blocked with complete media/10% (v/v) human AB serum for 2 hours at room temperature.

Plates were incubated for 21 hours at 37°C, 5% CO_2 prior to spot development. Cells were discarded and plates washed five times with 1X PBS (200 µL/well). Bound cytokines were detected using BAM-peptide tagged mAb 7-B6-1, WASP-peptide tagged mAb MT504 and biotin-labelled mAb 12G8 diluted according to the manufacturer's instructions and 100 µL added to each well for 2 hours of incubation at room temperature. Plates were then washed five times with 1X PBS (200 µL/well) before adding the fluorophore conjugates: anti-BAM-490,

streptavidin-550 or anti-WASP-640 (each diluted 1:200 in 1X PBS/0.1% [v/v] BSA), followed by a 1 hour incubation at room temperature. Plates were washed with 1X PBS (200 µL/well) and 50 µL/well of fluorescence enhancer added for a 15 minute incubation at room temperature. After discarding the fluorescence enhancer, the underdrain (plastic attached to the bottom surface of the plate) was removed before leaving the plates to dry in the dark. The numbers of spot forming units (SFU) were determined using a Mabtech IRIS FluoroSpot reader (Mabtech, Sweden) with software enabling overlay analysis of cells secreting all three cytokines. Stimulation index (SI) was calculated as follows: SI = total sum of SFUs in peptide-containing wells/total sum of SFU in negative control (DMSO) wells. To identify responders, Receiver Operating Characteristic (ROC) curve analysis performed previously using ELISpot data (Arif et al., 2004) indicated an optimal SI cut-off of 3.0 for IFN-γ. For consistency, a cut-off of 3.0 was also applied to IL-17A and IL-22.

2.5.3 Statistical analysis

The number of individuals eliciting a positive cytokine response was compared between different peptides using Fisher's exact test (QuickCalcs GraphPad). Analyses of paired SI IFN-γ responses to short and long hybrid peptides were carried out using the Wilcoxon signed rank test for pairing (GraphPad Prism). A 2x2 contingency table was also used to compare hybrid polypeptide responders and non-responders firstly, to that of the corresponding reversed orientation sequence and secondly, to the short epitope sequence and a p-value computed using Fisher's exact test. Differences were considered statistically significant at $p < 0.05$.

2.6 *In silico* data analysis

Peptides with a FDR less than 1% were discarded and duplicate peptides deleted. For binding motif prediction, the standard sequence logo option of the programme Icelogo (Colaert et al.,

2009) was used. Subcellular location of the human proteome was matched to source proteins of eluted peptides. Human proteome data was retrieved from the UniProt database (UniProt Consortium, 2012) using only reviewed records that have subcellular location annotations. For binding affinity prediction of eluted peptides and overlapping 15-mers the NetMHCII 2.3 algorithm (Jensen et al., 2018) was used. The NetMHCII 2.3 Server was also used for binding core sequence prediction. The program SitePrediction (Verspurten et al., 2009) was used to identify cathepsin cleavage sites in the C-peptide:WE14 sequence.

CHAPTER 3: OPTIMISATION OF AN ANTIGEN DELIVERY SYSTEM

3.1 Background to the Chapter

The hypothesis to be tested is that the hybrid peptides that have been described (Baker et al., 2018; Baker et al., 2019b; Delong et al., 2016; Wiles et al., 2019) derive from a larger species i.e., a hybrid polypeptide which requires immunological processing via canonical antigen presentation pathways. To address this hypothesis, I aim to identify novel naturally processed and presented epitopes (NPPEs) derived from a hybrid polypeptide, generated by the HLA class II processing pathway. In order to do so, the polypeptide antigen needs to be targeted to the surface of APCs for uptake, processing and presentation. This chapter explores the use of two methods of antigen delivery, pokeweed mitogen (PWM) and DEC205 (CD205) to optimise delivery of a long peptide containing a known epitope.

Initial attempts to study epitopes of exogenous antigens naturally presented by APCs adopted pulsing of EBV B cells with recombinant autoantigen in the context of Goodpasture disease (Phelps et al., 1996). Insufficient peptide material was obtained to identify epitopes bound to HLA-DR15, likely due to poor passive uptake by B cells (Chicz et al., 1992). Hence, the need for an antigen delivery system (ADS) is critical to overcome this issue. (Peakman et al., 1999) pioneered development of a lectin-based ADS to identify NPPEs derived from the intracellular domain of insulinoma-associated antigen-2 (IA-2ic) presented by HLA-DR4. Here the authors target antigen onto the surface of EBV B cells by incubation sequentially on ice with biotinylated pokeweed mitogen (bPWM), avidin-D and biotinylated IA-2ic. bPWM binds preferentially to carbohydrate moieties on cell surface receptors with immunoglobulin-like domains e.g. the BCR (Chilson and Kelly-Chilson, 1989). Avidin is then used as a bridge between bPWM bound to the cell surface and biotinylated IA-2ic permitting rapid internalisation, processing and presentation following incubation at 37°C. Synthetic peptides

of IA-2ic based on NPPEs identified by elution were found to bind HLA-DR4 and stimulate T cell proliferation in HLA-DR4-positive patients but not in controls. More recently, epitopes of GAD65 have also been eluted by pulsing of EBV B cells with PWM crosslinked to whole GAD65 (Peakman, unpublished). Presumably, efficient delivery of antigen in these systems is due to the ability of PWM to enhance APC function, because it is a polyclonal activator. Altogether, these findings support the use of a PWM-based ADS in directing antigens into the class II processing pathway and the generation of disease-relevant epitopes.

3.2 Use of DEC205 for APC targeting

DEC205 is a C-type multilectin transmembrane receptor expressed preferentially by DCs and thymic epithelial cells (Jiang et al., 1995). C-type lectin receptors are a diverse group of receptors originally identified as Ca^{2+}-dependent proteins capable of binding carbohydrate moieties of endogenous and exogenous origin to trigger phagocytic responses (Figdor et al., 2002). In DCs, the DEC205 receptor is endocytosed into clathrin-coated pits and recycled through late endosomal/lysosomal compartments rich in MHC class II molecules, leading to efficient antigen presentation to $CD4^+$ T cells (Mahnke et al., 2000). B cells express low levels of DEC205 but this increases significantly upon activation (Kamphorst et al., 2010). Thus, given its potent role in antigen presentation in DCs, DEC205-targeted antigen delivery in B cells has also gathered interest (Gurer et al., 2008; Leung et al., 2013).

Leung and colleagues (2013) were able to show that EBV-transformed B cells express similar levels of DEC205 to DCs and are capable of processing and presenting DEC205-targeted EBV antigens to peptide-specific $CD4^+$ and $CD8^+$ T cell clones. To target EBV antigens to the DEC205 receptor the authors cloned and purified fusion antibodies containing the COOH terminus of a monoclonal anti-human DEC205 antibody fused to cDNA encoding a sequence

of amino acids derived from different EBV antigens. Interestingly, incubation of anti-DEC205 fusion antibodies with EBV B cells stimulated higher IFN-γ secretion by T cells than targeting DEC205 on monocyte-derived DCs. The authors propose that EBV B cells retain antigen targeted by DEC205 for longer than DCs therefore prolonging T cell activation. Undoubtedly, targeting of DEC205 posed an additional method of antigen delivery to be investigated in the current study.

To first confirm expression of DEC205 on an EBV B cell line harbouring the T1D-associated HLA-DR4-DQ8 haplotype, WT51 cells were stained with an anti-DEC205 monoclonal antibody (αDEC205) and analysed by flow cytometry (Figure 3.1). Consistent with the findings of (Leung et al., 2013), DEC205 was expressed at high levels prompting further exploration as to whether this could be exploited in an ADS.

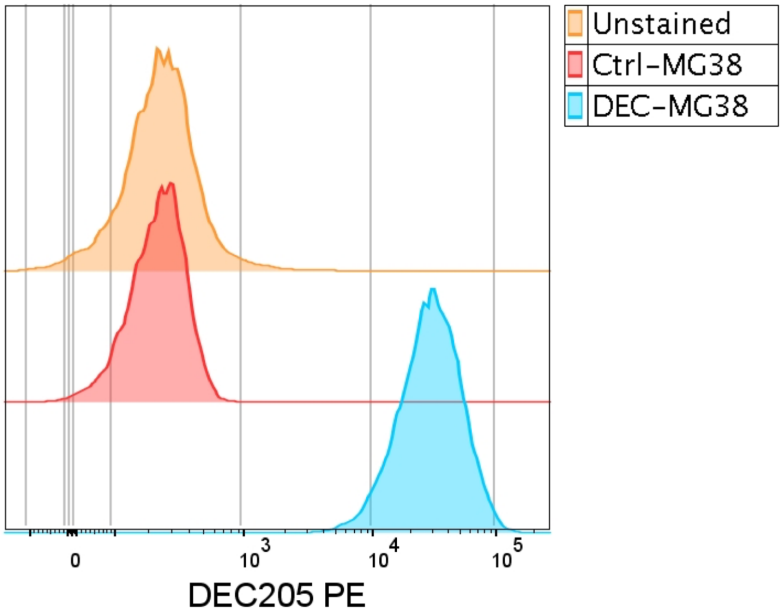


Figure 3.1 Cell surface expression of DEC205 on WT51 B cell line. WT51 cells were analysed for surface expression of DEC205 (blue histogram) by flow cytometry (red histogram; isotype control).

3.3 Construction of antigen delivery strategies

To evaluate PWM and DEC205 as putative delivery systems, optimisation studies focused on processing and presentation of a 41-mer peptide containing a HLA-DR4-restricted GAD65(274-286) epitope embedded within a hybrid polypeptide (control peptide in Table 2.1). The presence of a C-terminal sulfhydryl group (-SH) permits covalent conjugation of the 41-mer peptide to amine-containing (-NH_2) proteins such as PWM or an anti-human DEC205 antibody via a heterobifunctional crosslinker (Figure 3.2). PWM- and αDEC205-peptide conjugates were used to pulse WT51 cells for 30 minutes on ice at high cell density to increase fractional saturation (number of receptors occupied by ligand). To adjust the cell concentration to a suitable density for overnight culture at 37°C, fresh culture medium was added, thus diluting the peptide conjugate. T33.1 hybridoma cells (recognise the GAD65[274-286] epitope in the context of HLA-DR4) (Wicker et al., 1996) were used as a readout of antigen presentation by co-culture with pulsed B cells (Figure 3.3).

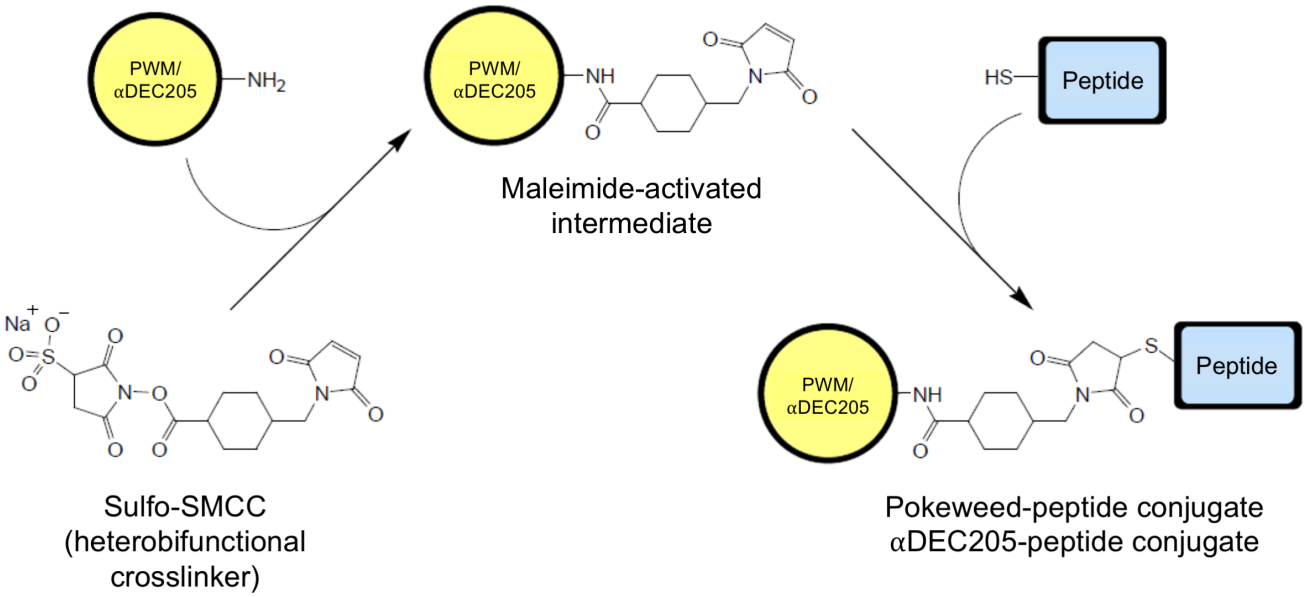


Figure 3.2 Crosslinking of peptide to carrier protein. Sulfo-SMCC activates amine groups within the carrier protein to form a maleimide-activated intermediate. This intermediate spontaneously reacts with free sulfhydryl groups of cysteine residues within the peptide to form a stable covalent linkage. Adapted from SMCC and sulfo-SMCC product instructions (ThermoFisher). PWM, pokeweed mitogen; αDEC205, anti-DEC205 antibody.

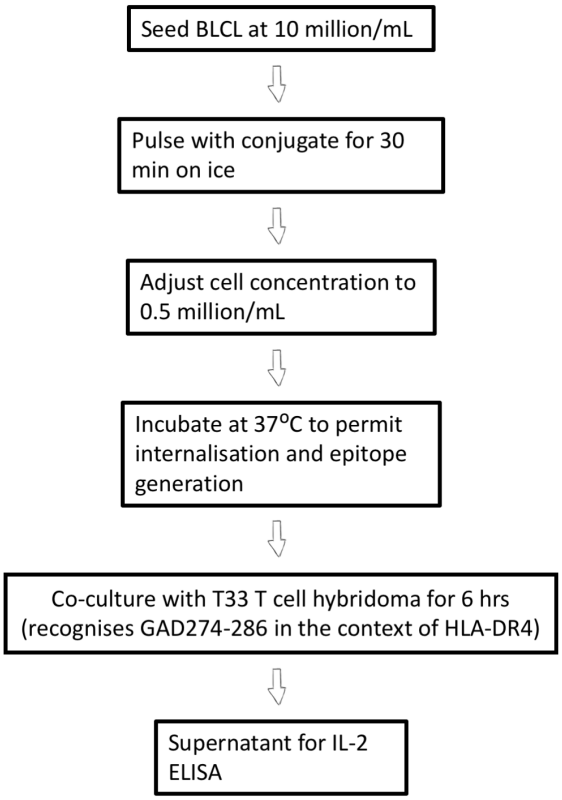


Figure 3.3 Schematic of ADS pulsing. As presentation of the GAD65(274-286) epitope is restricted to HLA-DR4, an EBV B lymphoblastoid cell line (BLCL) expressing HLA-DR4 was used for pulsing experiments.

3.4 Assessing ADS efficiency of antigen presentation

To be an efficient ADS, covalent conjugation to a carrier protein should enhance delivery of the GAD65(274-286)-containing 41-mer peptide. This was assessed by comparing activation of the T33 hybridoma by co-culture with either WT51 cells pulsed with conjugated peptide or co-incubated with 41-mer peptide for 16 hours in the absence of pulsing (Figure 3.4A). As a positive control, prior to co-culture WT51 cells were pulsed with the short 13-mer GAD65(274-286) epitope for 30 minutes at 37°C at equimolar concentration to the 41-mer peptide. The GAD65(274-286) epitope binds externally to HLA-DR4 molecules present on the APC therefore directly stimulating the T33 hybridoma. IL-2 production by T33 cells during co-culture with B cells pulsed with 41-mer peptide or activated PWM alone was also measured. In the case of activated PWM, addition of the crosslinker to crude PWM has generated a stable maleimide-activated intermediate but deprived of peptide for conjugation.

The data show that co-incubation of B cells with a titration of 41-mer peptide for 16 hours induced dose-dependent IL-2 secretion by the T cell hybridoma (Figure 3.4B, C). This indicated that the GAD65(274-286) epitope can be generated from the 41-mer peptide and passive uptake by the WT51 B cell line is possible. Pulsing with 2.18 μM of PWM-conjugated peptide for 30 minutes followed by twenty-fold dilution of the peptide (0.1 μM final), results in a similar level of T cell activation (approximately 60%, normalised to the GAD65[274-286] positive control) to co-incubation with 2.18 μM of 41-mer peptide for 16 hours. Similar T cell activation (3-5%) was also observed following pulsing with 0.218 μM of PWM-peptide conjugate (0.01 μM final) and co-incubation with 0.218 μM of 41-mer peptide, although IL-2 levels appear indistinguishable from the activated PWM alone control. Moreover, B cells pulsed with the 41-mer peptide alone did not stimulate the T33 hybridoma and co-culture with activated PWM-pulsed B cells produced a low background level of IL-2. These data indicate that the 41-mer peptide lacks the capacity to bind directly to HLA-DR4 present on the APC surface and then present GAD65(274-286); and that activated PWM does not have a direct mitogenic effect on the T33 hybridoma. Thus, individually the components within the PWM-peptide conjugate do not cause activation of the T cell hybridoma.

In further support of the need for conjugation, pulsing of WT51 B cells with non-conjugated activated PWM and 41-mer peptide stimulates IL-2 secretion above background (7% of GAD65[274-286] response) but this does not reach the same level of T cell activation achieved if the two components are conjugated (~60%). Most likely this response is due to spontaneous conjugation of the two components, even at 4°C, albeit at a low level. Surprisingly, pulsing with an αDEC205-peptide conjugate yielded no T cell response at either of the concentrations tested. Taken together, these findings suggest that twenty-fold more peptide is necessary for 16 hours' incubation to reach a similar T cell activation compared with a PWM-based ADS,

demonstrating enhancement in the delivery of the 41-mer peptide into the APC for processing. All downstream pulsing experiments were performed using the peptide molarity shown in Figure 3.4 to give the greatest level of T cell activation (2.18 μM during 30 min pulse, 0.1 μM final).

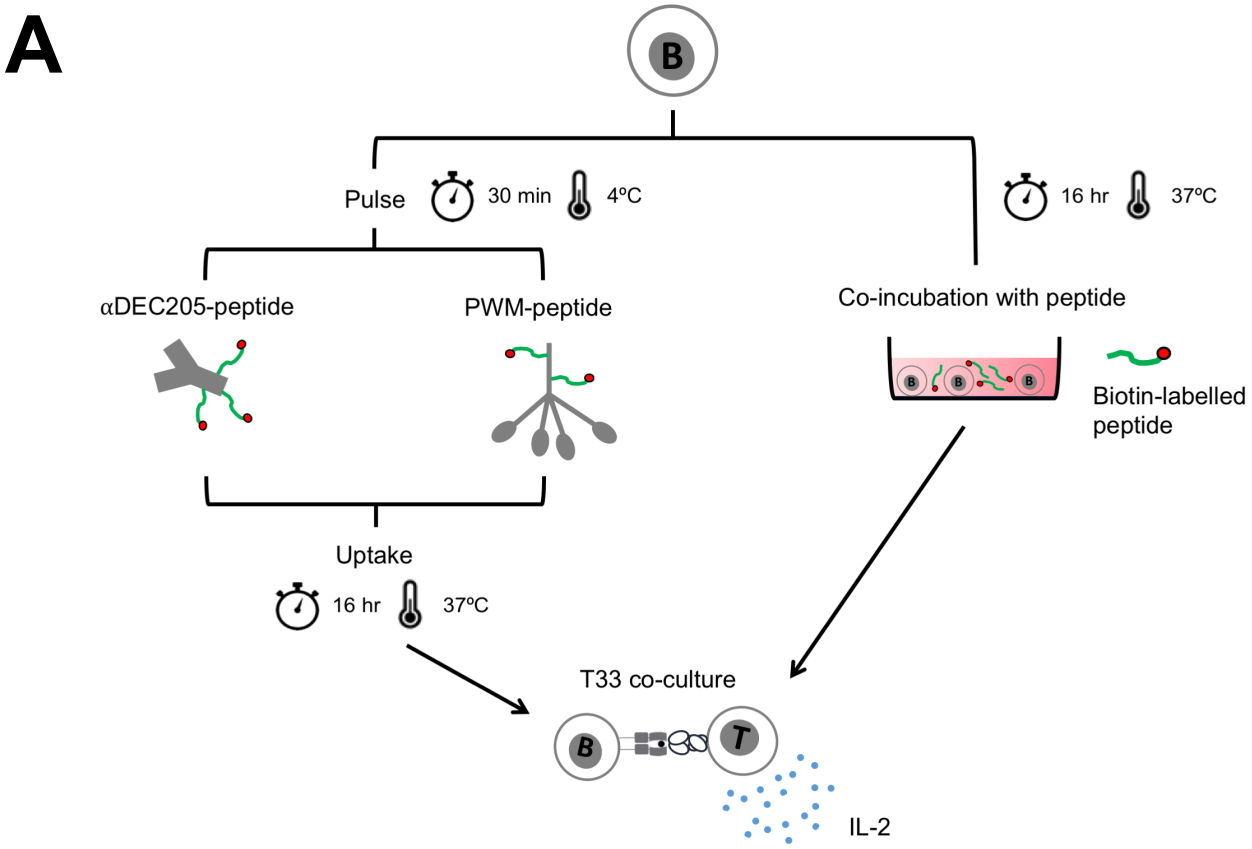


Figure 3.4 Pulsing with PWM-peptide conjugate enhances delivery of a 41-mer peptide containing an immunodominant T cell epitope from GAD65. A depicts a schematic of the experimental strategy. IL-2 production by the GAD65(274-286)-specific T33 hybridoma was measured following co-culture with WT51 EBV B cells (APC) either co-incubated with a 41-mer GAD65(274-286)-containing peptide for 16 hours or pulsed with an ADS. Within the ADS, the 41-mer peptide is conjugated to either PWM or αDEC205. A hyphen signifies conjugation of ADS components. B and C show data from a representative experiment and combined data from three independent experiments expressed as percentage of stimulation with GAD65(274-286) peptide (positive control), respectively. For experiments involving pulsing with conjugate, the final molarity of 41-mer peptide in the culture medium for overnight incubation at 37°C is indicated in brackets. Graphs indicate mean ± SD as well as individual data points representing either a separate T cell-B cell co-culture (B) or an independent experiment (C).

B

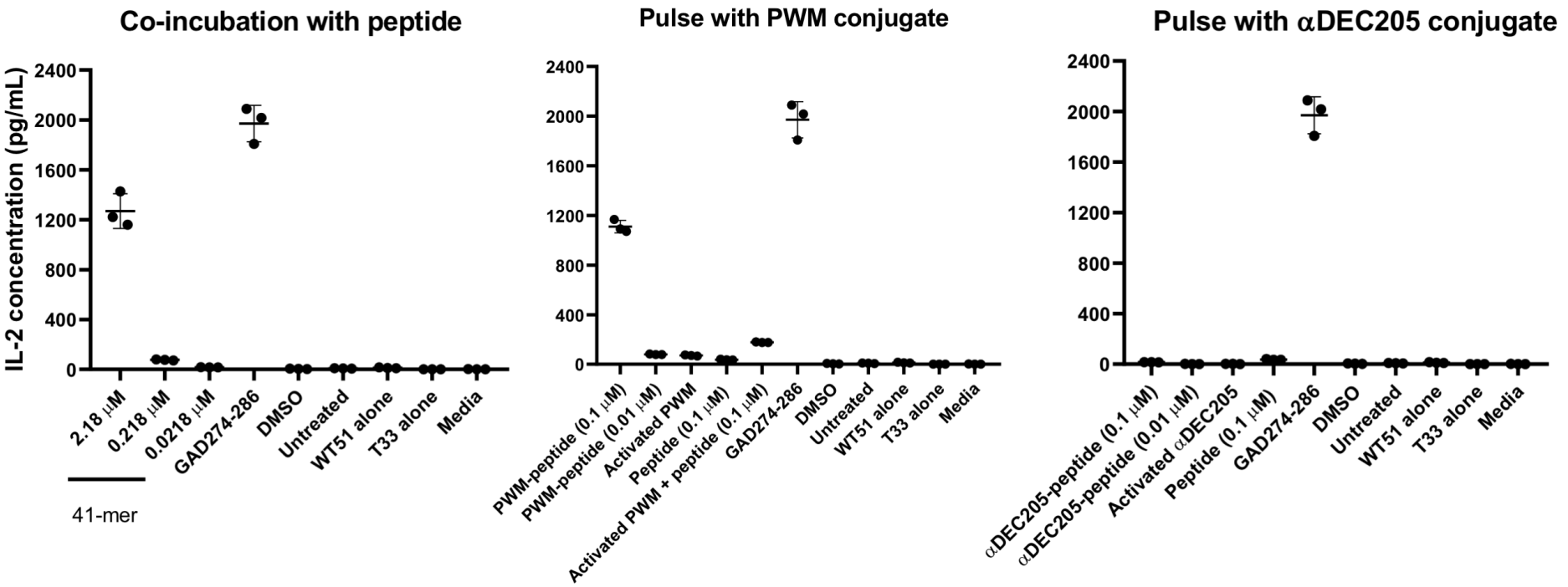


Figure 3.4 Pulsing with PWM-peptide conjugate enhances delivery of a 41-mer peptide containing an immunodominant T cell epitope from GAD65. A depicts a schematic of the experimental strategy. IL-2 production by the GAD65(274-286)-specific T33 hybridoma was measured following co-culture with WT51 EBV B cells (APC) either co-incubated with a 41-mer GAD65(274-286)-containing peptide for 16 hours or pulsed with an ADS. Within the ADS, the 41-mer peptide is conjugated to either PWM or αDEC205. A hyphen signifies conjugation of ADS components. B and C show data from a representative experiment and combined data from three independent experiments expressed as percentage of stimulation with GAD65(274-286) peptide (positive control), respectively. For experiments involving pulsing with conjugate, the final molarity of 41-mer peptide in the culture medium for overnight incubation at 37°C is indicated in brackets. Graphs indicate mean ± SD as well as individual data points representing either a separate T cell-B cell co-culture (B) or an independent experiment (C).

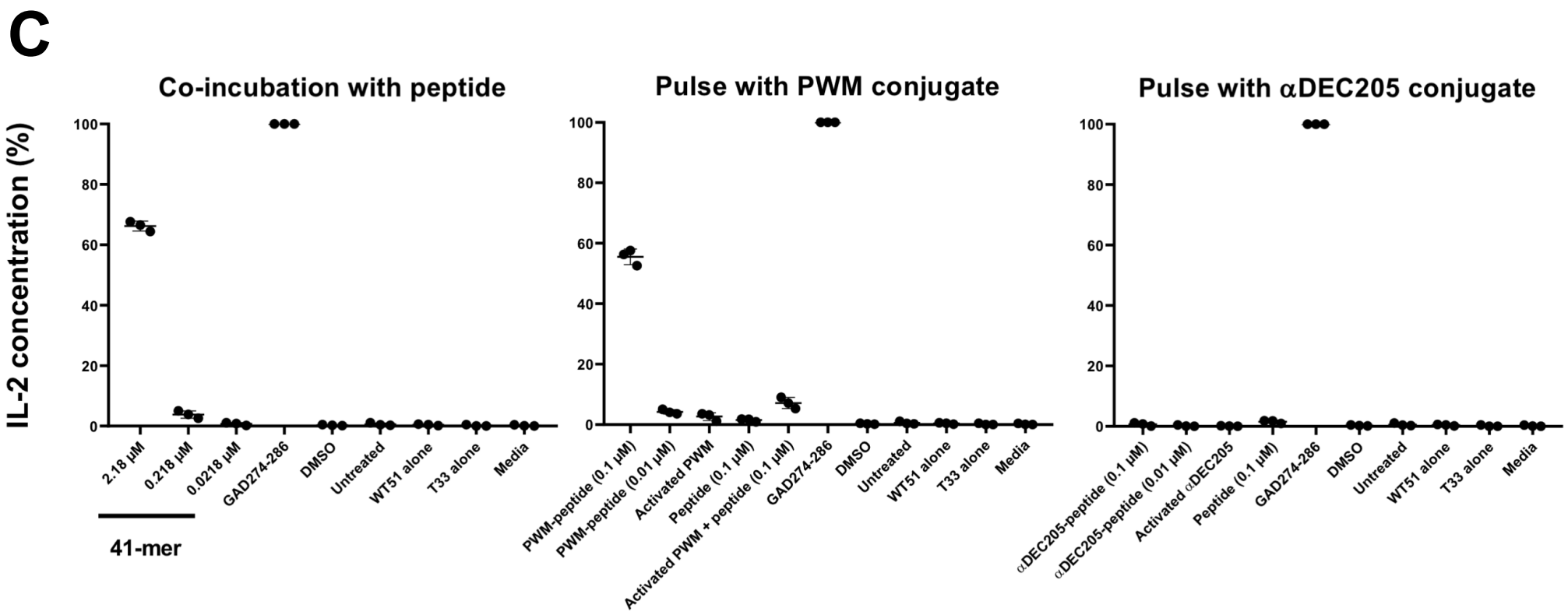


Figure 3.4 Pulsing with PWM-peptide conjugate enhances delivery of a 41-mer peptide containing an immunodominant T cell epitope from GAD65. A depicts a schematic of the experimental strategy. IL-2 production by the GAD65(274-286)-specific T33 hybridoma was measured following co-culture with WT51 EBV B cells (APC) either co-incubated with a 41-mer GAD65(274-286)-containing peptide for 16 hours or pulsed with an ADS. Within the ADS, the 41-mer peptide is conjugated to either PWM or αDEC205. A hyphen signifies conjugation of ADS components. B and C show data from a representative experiment and combined data from three independent experiments expressed as percentage of stimulation with GAD65(274-286) peptide (positive control), respectively. For experiments involving pulsing with conjugate, the final molarity of 41-mer peptide in the culture medium for overnight incubation at 37°C is indicated in brackets. Graphs indicate mean ± SD as well as individual data points representing either a separate T cell-B cell co-culture (B) or an independent experiment (C).

3.5 HLA-restriction of presented epitopes derived from ADS

To eliminate the possibility that the response to the PWM-peptide conjugate occurs independently of the presence of HLA-DR4, T33 hybridoma cells were co-cultured with pulsed EBV B cells expressing either HLA-DR4 or HLA-DR3 (Figure 3.5A). Both B cell lines were also pulsed with activated PWM, peptide alone or prior to T cell-B cell co-culture, the short GAD65(274-286) epitope. Since preceding experiments (Figure 3.4) established that the response is due to the peptide and not DMSO (peptide vehicle), non-pulsed cells were used as a control. Consistent with previous findings, pulsing of a HLA-DR4+ EBV B cell line with PWM-peptide conjugate followed by T33 co-culture provoked significant IL-2 secretion (>2000 pg/mL). The amount of IL-2 however, was variable between two independent experiments (Figure 3.5B, C) due to inherent variability in the ability of the T33 hybridoma to respond between experiments. Importantly, this IL-2 response is abolished upon co-culture of the T33 hybridoma with a HLA-DR3-expressing B cell line pulsed with the PWM-based ADS. B cells pulsed with the 41-mer peptide alone did not activate the T33 hybridoma above background regardless of the HLA type of the B cell. Interestingly, pulsing of the HLA-DR3+ cell line with activated PWM produced a higher IL-2 response in both experiments than pulsing with the PWM-peptide conjugate, suggesting that the presence of the peptide may interfere with binding of the lectin.

A T cell response to the short GAD65(274-286) epitope was only observed by co-culture with an EBV B cell line expressing the HLA-DR4 haplotype. The response to the short epitope was lower in both experiments than pulsing with the conjugate; inconsistent with prior experiments. The likely explanation for this lies in the fact that the proportion of peptide within the conjugate that is conjugated to PWM is unknown and may vary between batches. Overall, these data indicate that the response to the PWM-peptide conjugate is HLA-restricted but also peptide-

specific, as the peptide repertoire that is selected and presented by the HLA-DR3+ line will be different to that of the HLA-DR4+ line driven largely by differing anchor residues between different class II molecules (Bondinas et al., 2007).

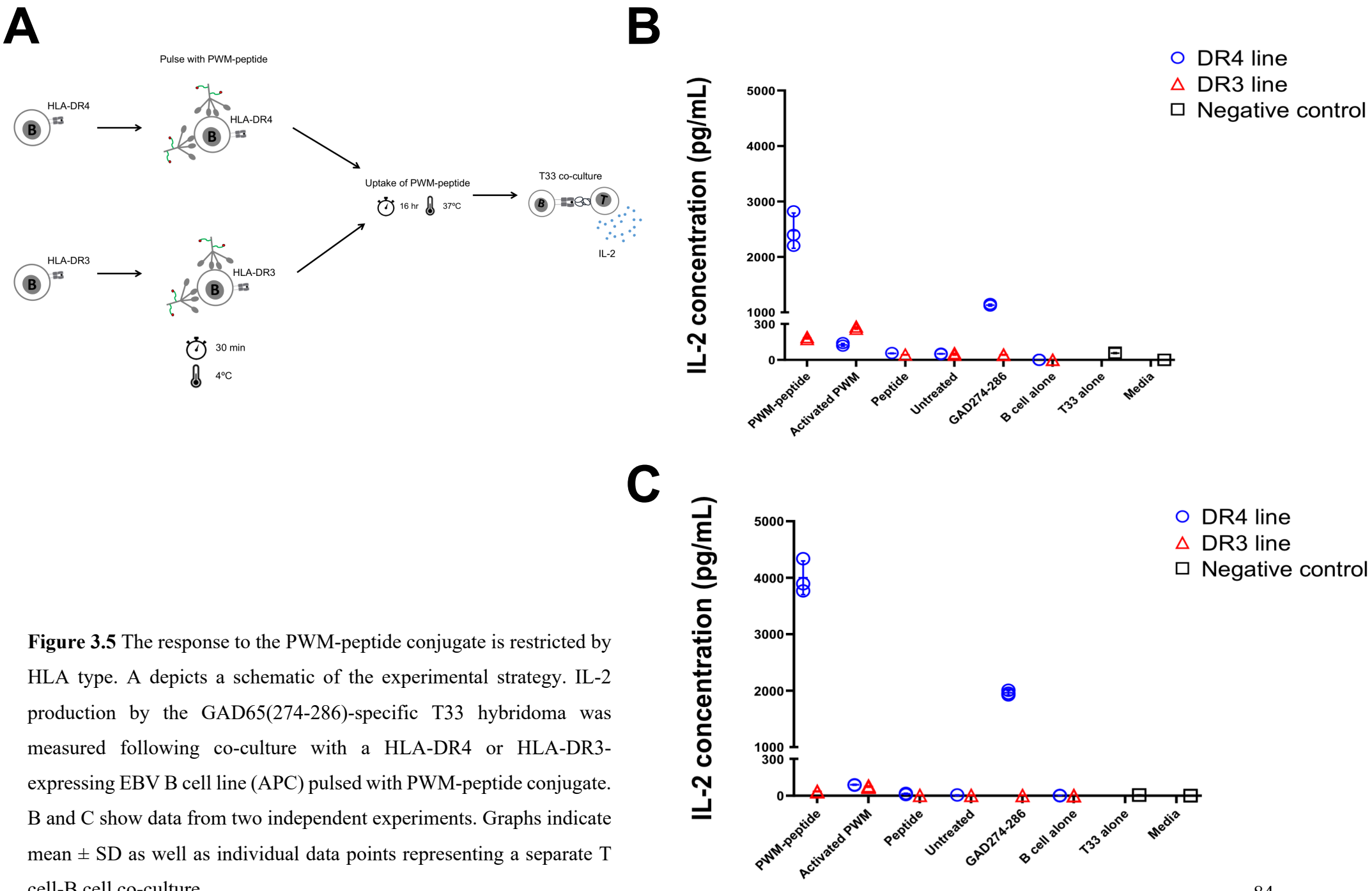


Figure 3.5 The response to the PWM-peptide conjugate is restricted by HLA type. A depicts a schematic of the experimental strategy. IL-2 production by the GAD65(274-286)-specific T33 hybridoma was measured following co-culture with a HLA-DR4 or HLA-DR3-expressing EBV B cell line (APC) pulsed with PWM-peptide conjugate. B and C show data from two independent experiments. Graphs indicate mean ± SD as well as individual data points representing a separate T cell-B cell co-culture.

3.6 Internalisation of labelled PWM

It was necessary to provide evidence that the PWM-peptide conjugate was being internalised rather than undergoing cell surface processing, reported to occur in the case of a fragment of tetanus toxin (Moss et al., 2007). Initial experiments to infer internalisation utilised the biotin moiety attached to the N-terminus of the 41-mer peptide. In this way, an anti-biotin antibody can be used to monitor the presence of 41-mer peptide at the surface of the APC by FACS. EBV B cells were pulsed with PWM-peptide conjugate, activated PWM or peptide alone and incubated at 37°C for 0, 1, 3 and 6 hours before placing the cells on ice (to stop further internalisation) followed by anti-biotin staining (Figure 3.6A).

Pulsing with activated PWM or peptide alone revealed a staining pattern similar to the isotype control, demonstrating an absence of peptide at the cell surface (Figure 3.6B). This further supports a lack of direct binding of the 41-mer peptide. The staining pattern of cells pulsed with PWM-peptide conjugate shows a gradual shift from 0 to 6 hours, indicated by a reduction in biotin median fluorescence intensity (MFI) (Figure 3.6C). Additionally, there appears to be a reduction in the number of positive events between 0 and 6 hours. The rate of reduction in biotin MFI was greatest between 0 and 1 hour, stabilising between 1 and 3 hours preceding further decline between 3 and 6 hours. Taken together, these results show that peptide is present at the cell surface after pulsing with PWM-peptide conjugate and decreases over time suggesting internalisation, particularly within one hour of pulsing.

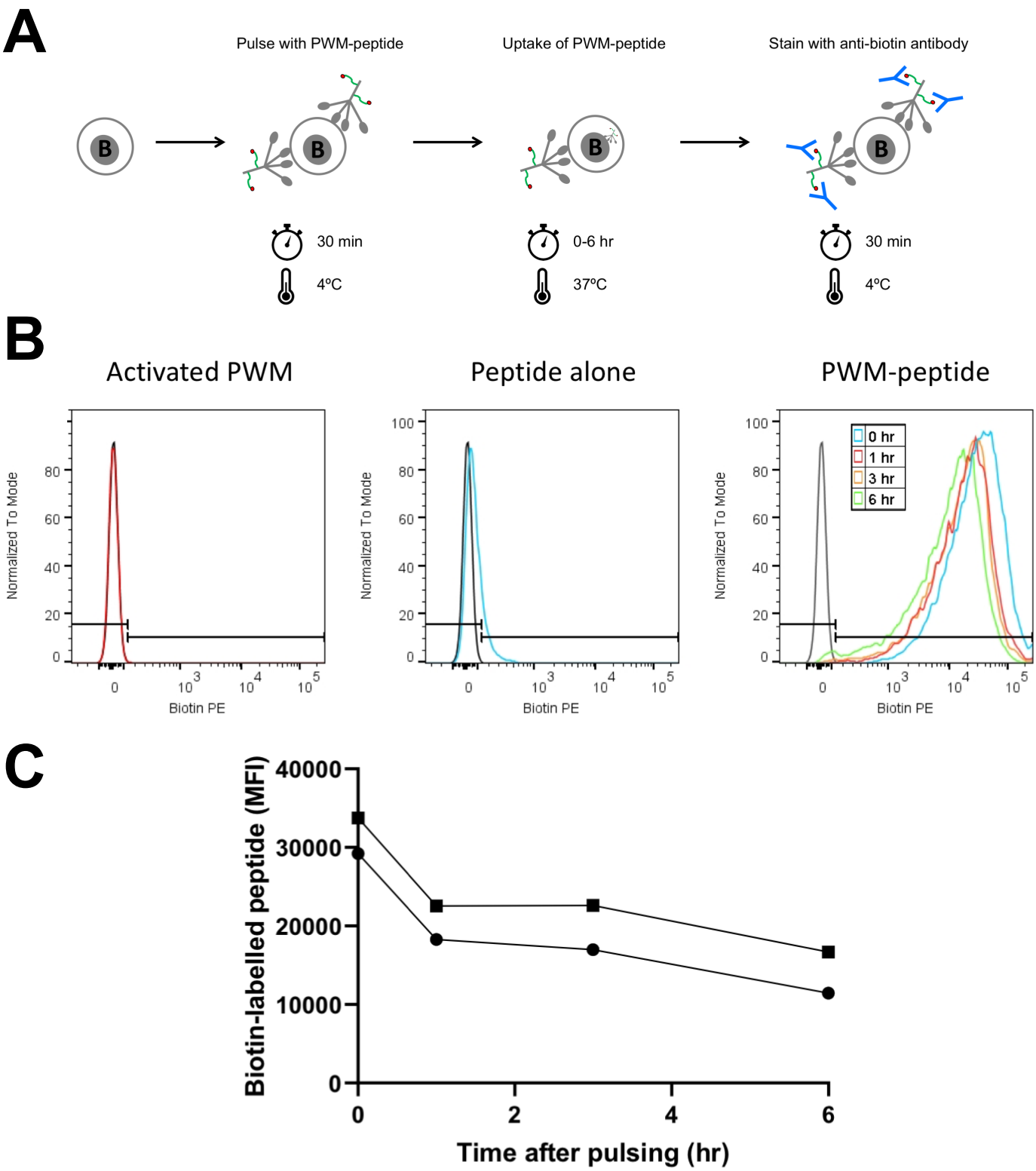


Figure 3.6 Biotinylated peptide is targeted to the cell surface after pulsing and disappears over time. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with activated PWM, peptide alone or PWM-peptide conjugate and stained with an anti-biotin antibody at specific time points following incubation at 37°C. B shows representative flow cytometry histograms illustrating the biotin staining pattern with black histograms denoting isotype control-stained cells. Biotin MFI values across the four time points are plotted from two independent experiments in C.

Although from the anti-biotin staining presented in Figure 3.6 it is compelling to speculate internalisation is occurring, an alternative explanation is that the peptide is simply falling off the surface of the cell over time. To confirm that this was not the case, imaging flow cytometry (ImageStream) was employed to visualise and quantify internalisation. In order to assess internalisation, PWM was conjugated to an allophycocyanin fluorescent label (PWM-APC) via available amine groups, thus replicating the same chemistry utilised by the sulfo-SMCC crosslinker in conjugating PWM to the 41-mer peptide. EBV B cells were pulsed with PWM-APC, incubated at 37°C and localisation of the fluorescent PWM compared at 0 and 6 hours (Figure 3.7A). The gating strategy applied for the ImageStream analysis is shown in Figure 2.2 (Materials and Methods).

To measure whether the localisation of PWM-APC was intracellular or surface-bound, a mask was designed that excludes the cell membrane and an internalisation score computed, reflecting the ratio of the amount of fluorescence located in the mask versus the total amount of fluorescence. Due to the focal depth, fluorescence from the top and the bottom of the cell will be out of focus thus contributing very little to the overall fluorescence. Internalisation scores around 0 have a mix of internalisation and membrane intensity. The higher the score, the greater the concentration of intensity inside the cell, as previously reported (Garcia-Vallejo et al., 2015). At 0 hours after pulsing with PWM-APC, the majority of cells present with a uniform staining pattern in a ring around the cell (Figure 3.7D, left) indicative of localisation at the plasma membrane. Because there is a degree of overlap in the fluorescence signal between inside and outside the cell, a median internalisation score of 0.3178 was observed at 0 hours. In contrast, after 6 hours of incubation at 37°C, a greater frequency of cells with significant intracellular staining are observable (Figure 3.7D, right). This is reflected by an increase in the median internalisation score from 0.3178 to 1.269 between 0 and 6 hours (Figure 3.7B, C).

These findings indicate a greater concentration of PWM-APC intensity inside the cell at 6 hours relative to 0 hours, demonstrating internalisation of fluorescently-labelled PWM.

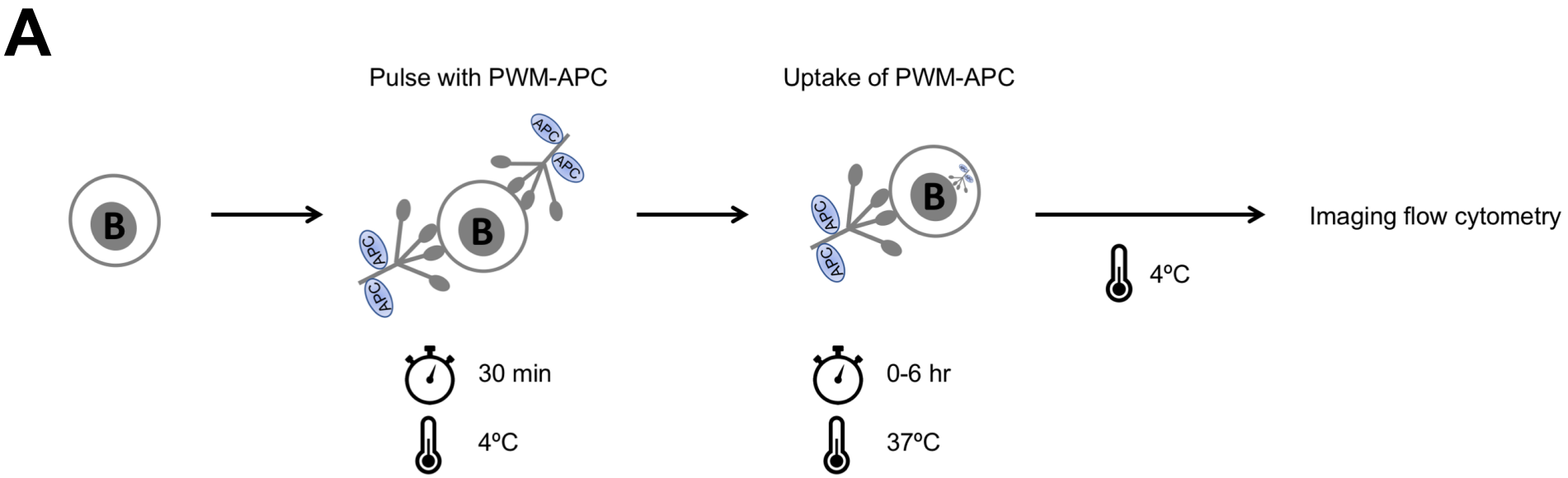
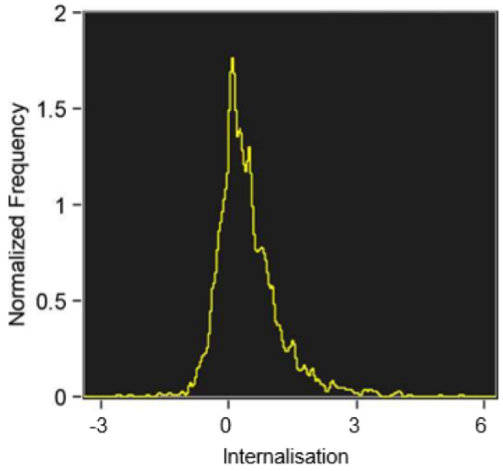


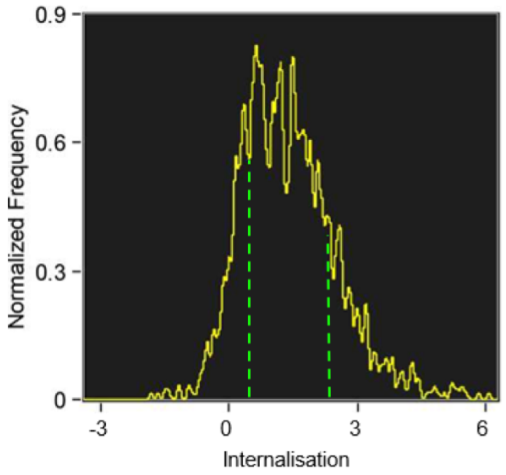
Figure 3.7 Internalisation of PWM-APC. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with PWM-APC and the degree of internalisation measured using ImageStream X. Single focused and cells positive for APC were gated. Internalisation scores were calculated, reflecting the ratio of cytoplasmic to total brightness intensity. B and C show histogram plots of internalisation scores for cells immediately after pulsing and after incubation for 6 hours at 37°C, respectively. Representative cell images captured by the Amnis ImageStream Flow Cytometer are also provided (D). Selected cell images were chosen to be well above and below the median with the first column showing brightfield images (Ch01) and the second column showing images of fluorescence of PWM (Ch11). Dotted green lines indicated in C correspond to the location on the histogram where representative images were chosen.

B



Population	Mean	Median
Single cells, focus & APC+	0.4588	0.3178

C



Population	Mean	Median
Single cells, focus & APC+	1.413	1.269

D

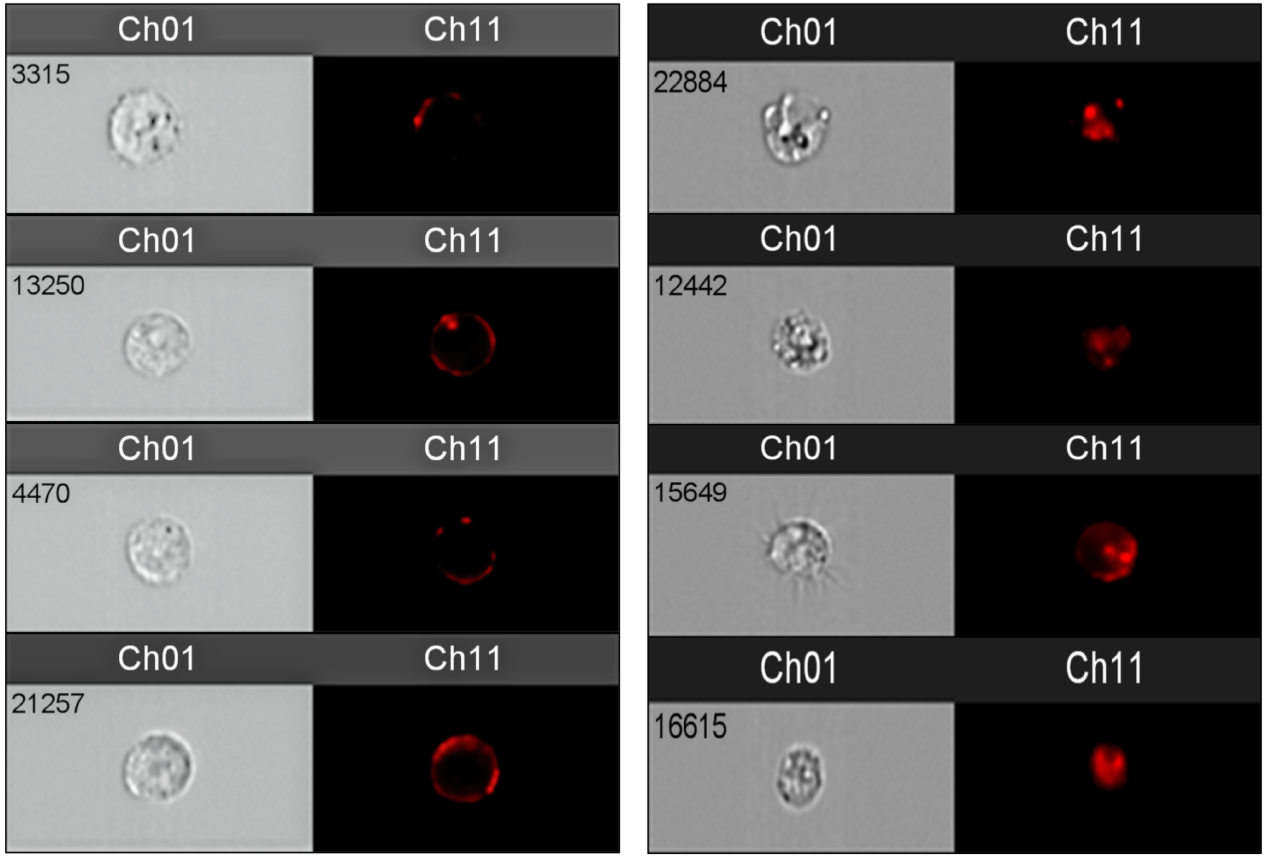


Figure 3.7 Internalisation of PWM-APC. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with PWM-APC and the degree of internalisation measured using ImageStream X. Single focused and cells positive for APC were gated. Internalisation scores were calculated, reflecting the ratio of cytoplasmic to total brightness intensity. B and C show histogram plots of internalisation scores for cells immediately after pulsing and after incubation for 6 hours at 37°C, respectively. Representative cell images captured by the Amnis ImageStream Flow Cytometer are also provided (D). Selected cell images were chosen to be well above and below the median with the first column showing brightfield images (Ch01) and the second column showing images of fluorescence of PWM (Ch11). Dotted green lines indicated in C correspond to the location on the histogram where representative images were chosen.

3.7 Changes in epitope density following PWM-mediated peptide delivery

Identification of naturally processed epitopes requires putative epitopes to be present on the cell surface at a reasonable abundance. It was therefore important to establish the time point after pulsing which lead to the maximal epitope density. To determine this using the GAD65(274-286)-containing 41-mer peptide, HLA-DR4-expressing EBV B cells were pulsed with PWM-peptide conjugate and incubated at 37°C for 0-6 hours before gentle fixation with 0.1% paraformaldehyde (Figure 3.8A). The fixation procedure terminates further processing by the B cell and co-culture with the T33 hybridoma provides an indication of the number of epitopes present at the cell surface at a given time point. In parallel, EBV B cells were also pulsed with activated PWM and fixed over the same time course as it is plausible to speculate that the response to activated PWM is reduced over time as PWM is internalised, hence providing less stimulation.

The data show a gradual increase in the number of epitopes present at the cell surface, indicated by an increase in IL-2 secretion, every 2 hours for the first 6 hours after pulsing with PWM-peptide conjugate (Figure 3.8B, C). This increase in T cell activation over time also provides additional evidence of processing and presentation which is known to be a highly dynamic process (Croft et al., 2013). On the other hand, pulsing with activated PWM and fixation at the various time points results in a level of IL-2 secretion which remains relatively unchanged throughout the time course, supporting the peptide-specific nature of the response to the conjugate. Interestingly, IL-2 secretion was observed by co-culture with B cells fixed immediately after pulsing with PWM-peptide conjugate. Given that the amount of IL-2 produced in this case (200 pg/mL) surpassed that of the activated PWM alone, it is unlikely that the response at 0 hours is due to low background stimulation by PWM. It is more likely that the fixation procedure is not completely efficient in rendering the cell incapable of processing.

Furthermore, one of the ligands PWM binds is the BCR, at which site internalisation has been described to commence within minutes after removal from ice (Hou et al., 2006). Thus, amongst the incubation steps there is the potential for epitope generation to occur at an early stage.

Of note, production of IL-2 by the T33 hybridoma is greatest upon co-culture with non-fixed B cells pulsed with PWM-peptide conjugate. This indicates the potential for a greater level of epitope density to be reached beyond 6 hours. For this reason, the time course was extended, incubating pulsed cells for 12-18 hours at 37°C before gentle fixation and co-culture with T33 hybridoma cells. Consistent between two independent experiments, the consequence of fixation either 14 or 16 hours after pulsing is a similar level of T cell activation (Figure 3.9A, B). A trend towards greater T cell activation is observed following fixation at 12 hours relative to 14-16 hours; however, the error bars rule overlap suggesting there is not a significant difference between the conditions. Overall, the data suggest that the epitope density on the cell surface remains relatively constant 12-18 hours after pulsing with the PWM-based delivery system. To ensure feasibility of downstream pulsing experiments involving hybrid polypeptides, 16 hours (overnight incubation) was selected as the time point with the optimal epitope density.

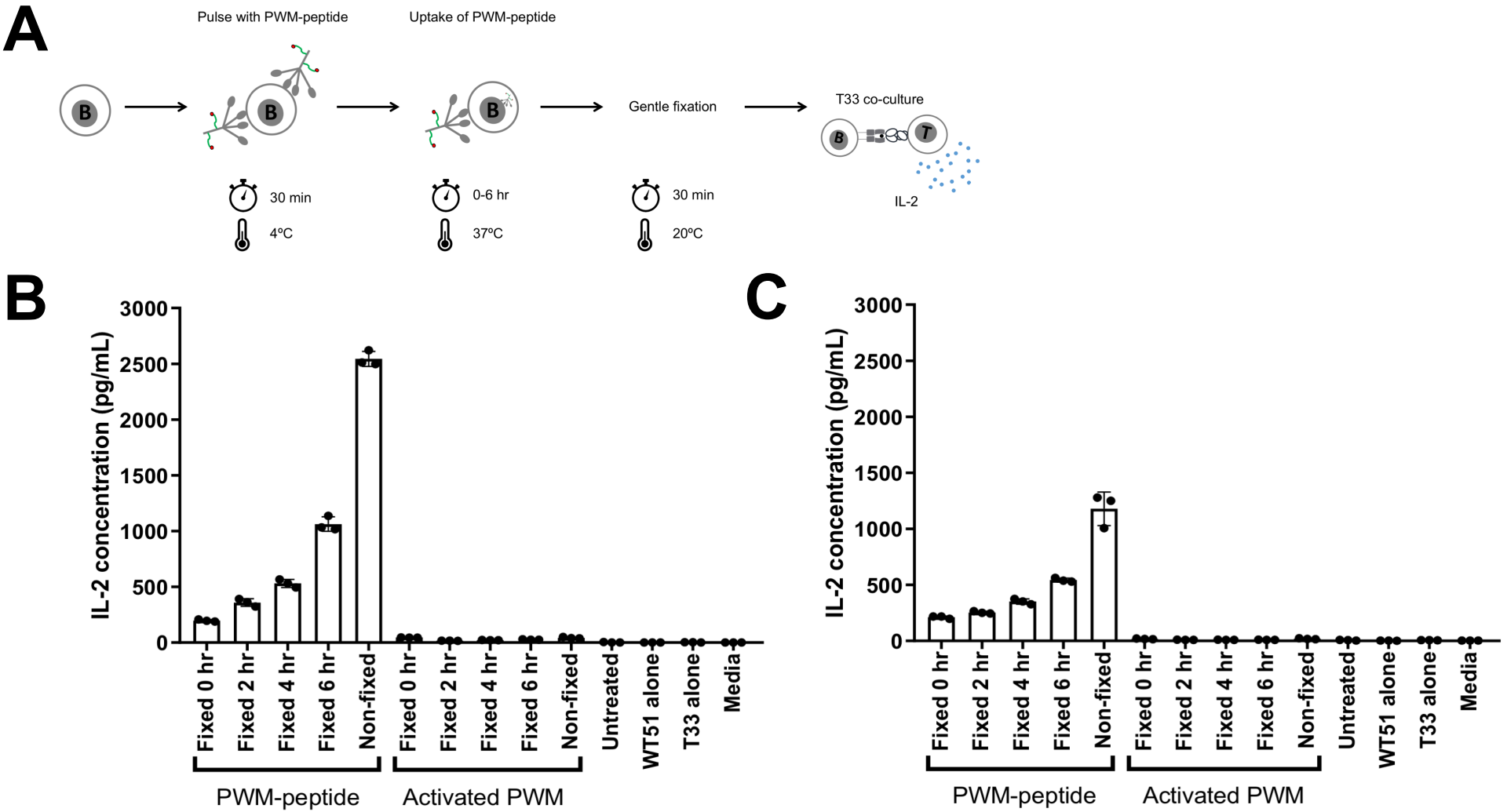


Figure 3.8 Epitope density on the cell surface increases 0-6 hrs after pulsing with PWM-peptide conjugate. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with either PWM-peptide conjugate or activated PWM and gently fixed every 2 hours. Fixed cells were co-cultured with the T33 hybridoma and IL-2 measured in the supernatant. B and C show data from two independent experiments. PMA and ionomycin was used as a positive control for IL-2 production in both experiments; mean values obtained for B and C were 60012 ± 7038 pg/mL and 37795 ± 1572 pg/mL, respectively. Graphs show mean ± SD as well as individual data points representing a separate T cell-B cell co-culture.

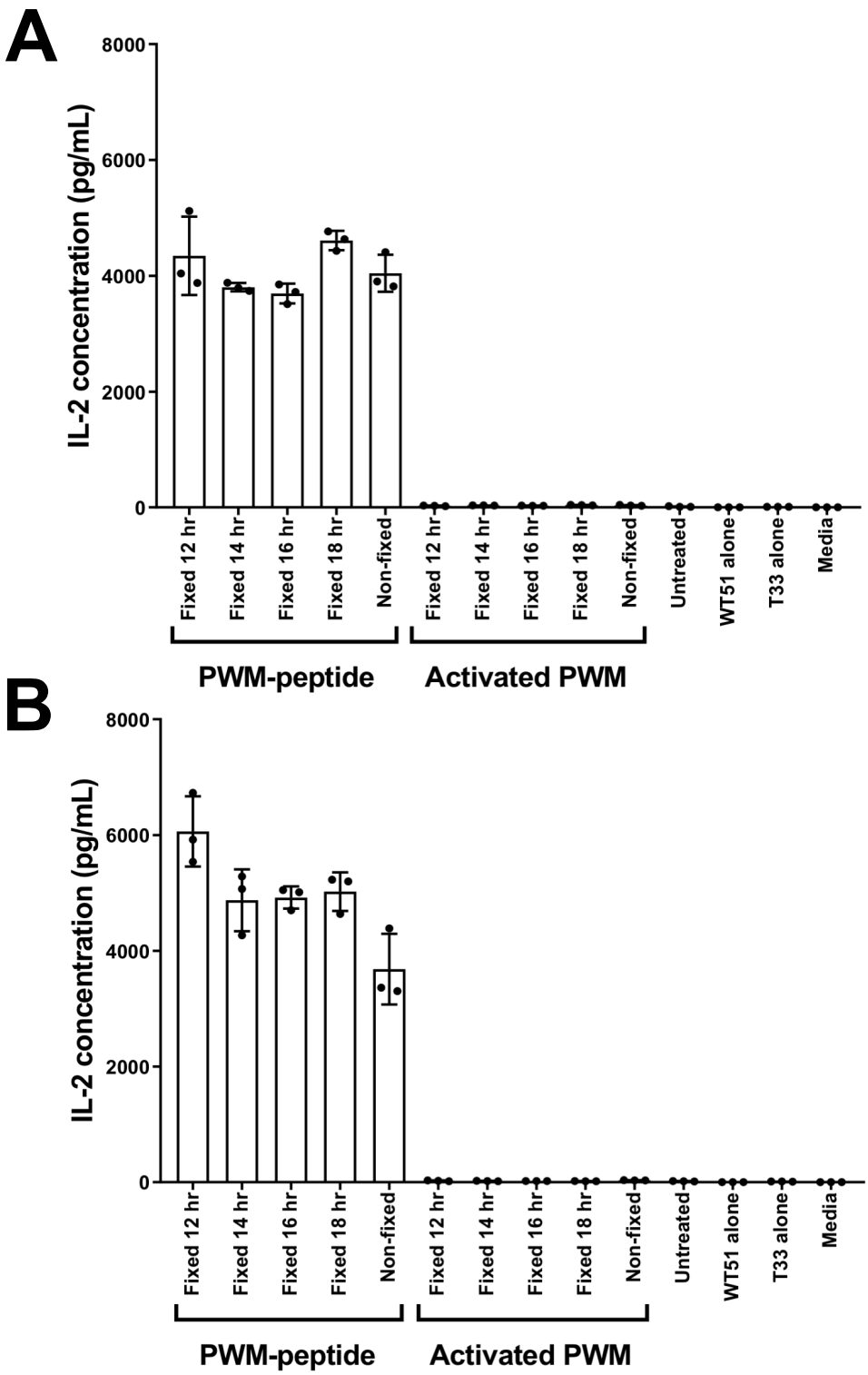


Figure 3.9 Epitope density on the cell surface remains constant 12-18 hrs after pulsing with PWM-peptide conjugate. EBV B cells were pulsed with either PWM-peptide conjugate or activated PWM and gently fixed at 2 hour intervals, commencing 12 hours after pulsing. Fixed cells were co-cultured with the T33 hybridoma and IL-2 measured in the supernatant. A and B show data from two independent experiments. PMA and ionomycin was used as a positive control for IL-2 production in both experiments; mean values obtained for A and B were 81710 ± 9377 pg/mL and 41281 ± 5451 pg/mL, respectively. Graphs show mean ± SD as well as individual data points representing a separate T cell-B cell co-culture.

3.8 Conjugation and cell surface binding of αDEC205-peptide conjugate

The lack of T cell response observed after pulsing with αDEC205-peptide conjugate is curious given a previous report using DEC205 to successfully target antigens for processing in EBV B cells (Leung et al., 2013). To explore this, it was first necessary to confirm conjugation of the 41-mer peptide to the anti-DEC205 antibody. Anti-mouse IgG beads bind any mouse kappa light chain-bearing immunoglobulin and were incubated with either activated αDEC205, 41-mer peptide or αDEC205-peptide conjugate (Figure 3.10A). To remove unbound antibody the beads were washed and subsequently incubated with streptavidin-APC. As the 41-mer peptide bears a biotin-label, providing the conjugate contains 41-mer peptide crosslinked to the antibody one would expect the beads to be APC-positive when analysed by flow cytometry. Indeed, as shown in Figure 3.10C, beads incubated with the conjugate stain positive for APC while beads incubated with the activated αDEC205 antibody do not due to the absence of 41-mer peptide. Beads incubated with the 41-mer peptide alone also stain negative for APC indicating that the 41-mer peptide does not bind to the anti-mouse IgG beads.

It was also necessary to exclude the possibility that conjugation to the 41-mer peptide disrupts the antigen-binding site of the antibody. This was investigated by pulsing WT51 cells with either αDEC205-peptide conjugate, activated αDEC205, non-activated αDEC205 or a negative control mouse IgG2b antibody. Pulsed cells were then stained with an anti-mouse IgG2b secondary antibody and analysed by flow cytometry (Figure 3.10B). Non-activated and activated αDEC205 show similar binding to the DEC205 antigen on the surface of the WT51 cells, while binding of the αDEC205-peptide conjugate is marginally reduced in comparison (Figure 3.10D). This reduction in binding was consistent across three independent experiments and carries the caveat that this could be due to the presence of the peptide interfering with the binding of αDEC205 to the cell surface. Alternatively, the presence of the peptide may interfere

with binding of the secondary antibody to αDEC205. Nevertheless, together these data demonstrate that biotinylated 41-mer peptide is detectable within the conjugate which still binds to the cell surface.

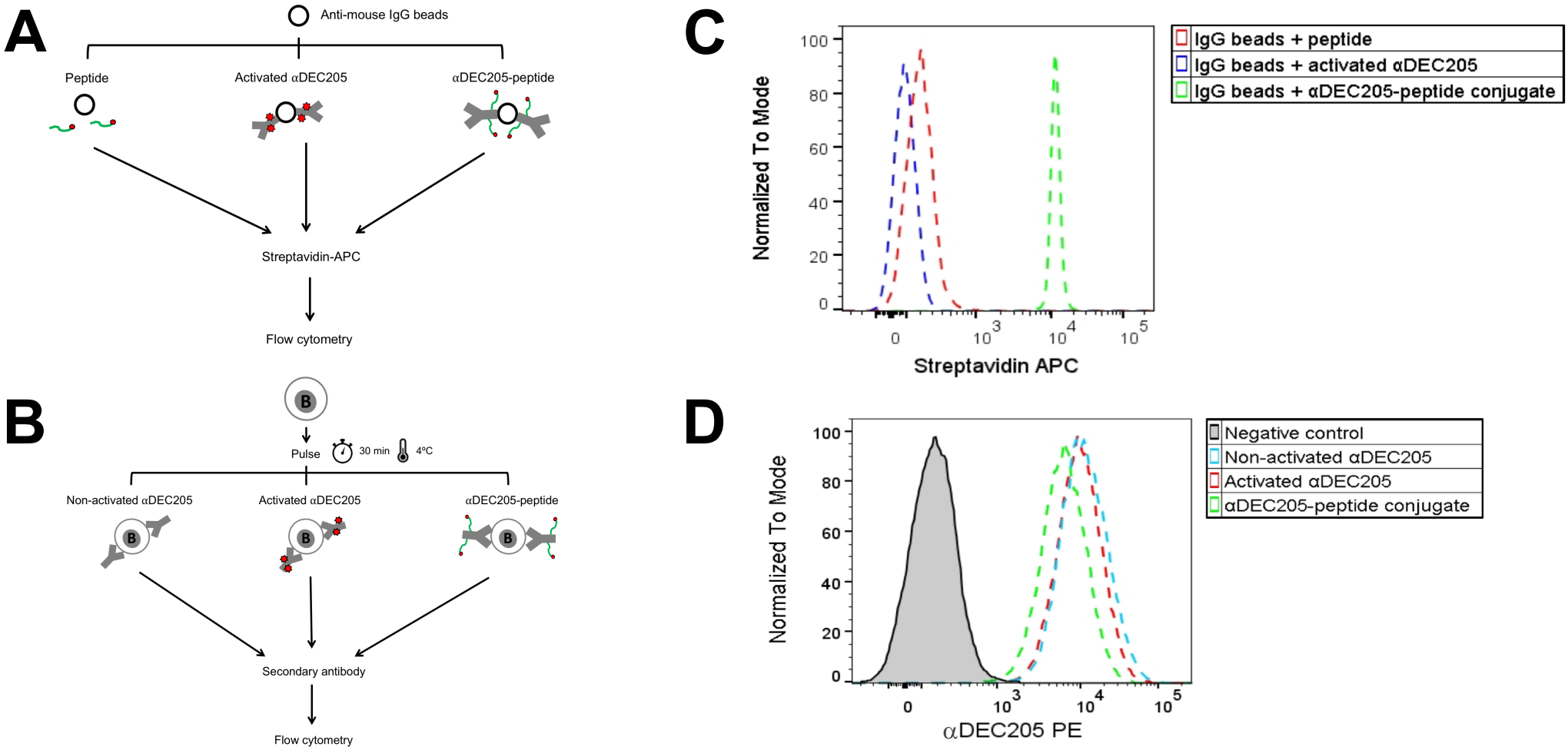


Figure 3.10 The αDEC205-peptide conjugate binds to the cell surface and contains biotinylated 41-mer peptide. A and B depict schematics of the experimental strategies. Anti-mouse IgG beads were incubated first with activated αDEC205, αDEC205-peptide conjugate or peptide alone and subsequently with streptavidin-APC. Beads were run on a flow cytometer and gated for APC staining (C). Histograms in D illustrate cells stained with an anti-mouse IgG2b secondary antibody after pulsing with non-activated αDEC205, activated αDEC205 or αDEC205-peptide conjugate to detect cell surface binding. Cells pulsed with a negative control mouse IgG2b antibody were also included (filled grey histogram). Data shown in C and D are representative of three independent experiments with similar results.

3.9 Internalisation of labelled anti-DEC205 monoclonal antibody

A failure of the αDEC205-peptide conjugate to be internalised would provide an explanation for the absence of a T cell response. To establish whether this was the case, EBV B cells were pulsed with either activated αDEC205, αDEC205-peptide conjugate or peptide alone and incubated at 37°C. At specific time points cells were removed from culture, placed on ice to stop further internalisation and stained with streptavidin-APC to detect the 41-mer peptide (biotin-labelled) on the APC surface (Figure 3.11A). Pulsing with activated αDEC205 resulted in a staining pattern identical to non-pulsed cells due to the absence of 41-mer peptide, whereas analysis of cells pulsed with peptide alone revealed a very low level of direct binding (Figure 3.11B). A striking reduction in the amount of biotin-labelled peptide present at the cell surface was observed between 0 and 1 hour post-pulse with the conjugate, falling to a level comparable to the staining seen with peptide alone. No further change in biotin MFI was observed beyond 1 hour (Figure 3.11C). These findings suggest rapid internalisation of the αDEC205-peptide conjugate within 1 hour at 37°C. However, this experiment merely suggests internalisation, therefore imaging flow cytometry was applied to be conclusive.

To examine internalisation of αDEC205 by imaging flow cytometry, a commercially available antibody labelled with phycoerythrin (αDEC205-PE) was used to pulse EBV B cells. Given that previous data suggest rapid internalisation of the αDEC205-peptide conjugate, cells were pulsed with αDEC205-PE and fluorescence localisation visualised immediately after pulsing and after 2 hours of incubation at 37°C (Figure 3.12A). In line with the ImageStream analysis performed with PWM-APC, the gating strategy outlined in Figure 2.2 (Materials and Methods) was applied but in this case PE-positive cells were gated. Calculation of median internalisation scores show an increase from 0.0939 at 0 hours to 1.244 at 2 hours, indicating that PE-labelled αDEC205 bound to the cell surface is being actively internalised over time (Figure 3.12B, C).

This can be visualised by a punctate staining pattern (Figure 3.12D, right) compared to the surface staining observed at 0 hours (Figure 3.12D, left). Although αDEC205 is labelled with phycoerythrin using unknown chemistry, the data shown in Figure 3.12 can be extrapolated to the αDEC205-peptide conjugate since I have shown binding of the antibody-peptide conjugate to the cell surface. It is therefore a reasonable assumption that the events which occur downstream of binding to the DEC205 antigen are the same between αDEC205-PE and the conjugate.

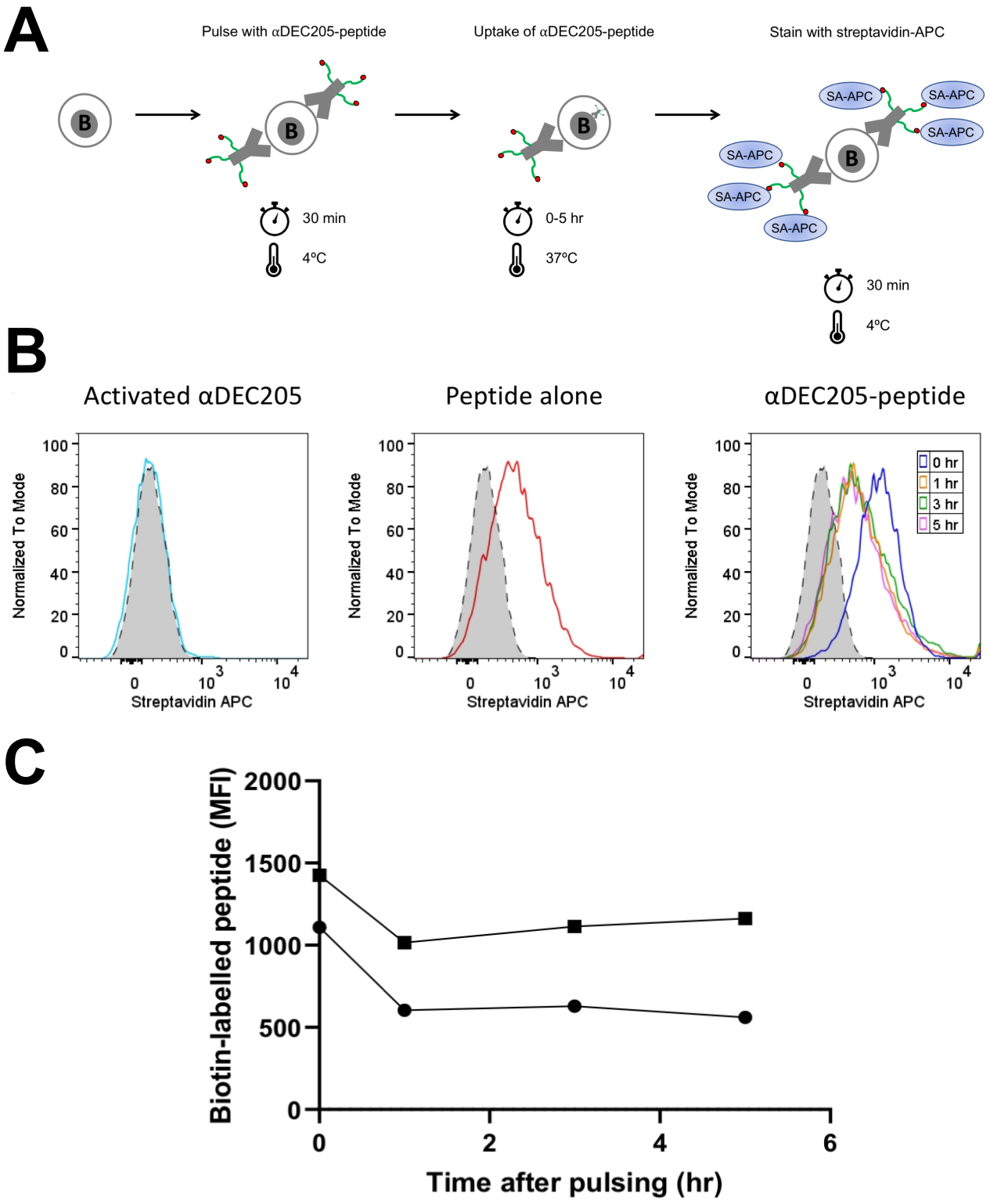


Figure 3.11 Reduction in the amount of peptide at the cell surface 1 hour after pulsing with αDEC205-peptide conjugate. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with activated αDEC205, peptide alone or αDEC205-peptide conjugate and stained with streptavidin-APC at specific time points following incubation at 37°C. B shows representative flow cytometry histograms illustrating the staining pattern with grey filled histograms denoting non-pulsed cells. Biotin MFI values across the four time points are plotted from two independent experiments in C.

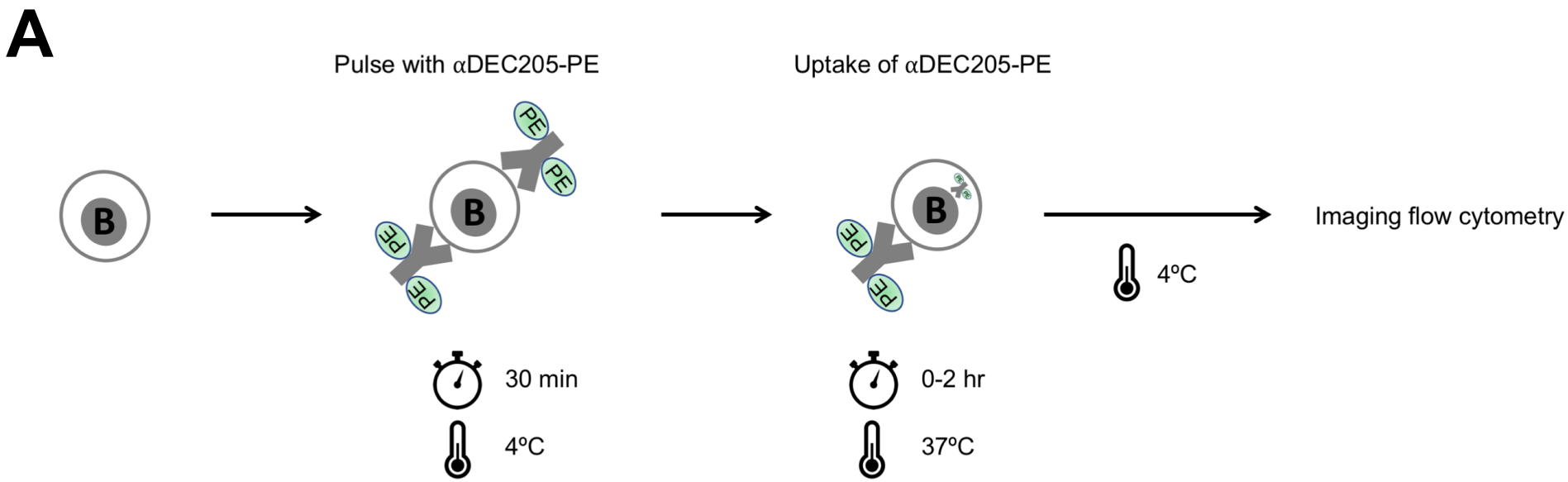
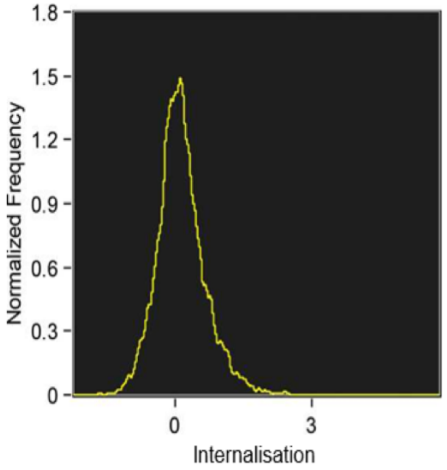


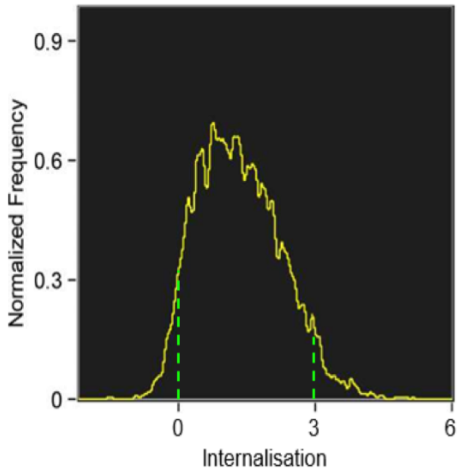
Figure 3.12 Internalisation of αDEC205-PE. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with αDEC205-PE and the degree of internalisation measured using ImageStream X. Single focused and cells positive for PE were gated. Internalisation scores were calculated, reflecting the ratio of cytoplasmic to total brightness intensity. B and C show histogram plots of internalisation scores for cells immediately after pulsing and after incubation for 2 hours at 37°C, respectively. Representative cell images captured by the Amnis ImageStream Flow Cytometer are also provided (D). Selected cell images were chosen to be well above and below the median with the first column showing brightfield images (Ch01) and the second column showing images of fluorescence of αDEC205 (Ch03). Dotted green lines indicated in C correspond to the location on the histogram where representative images were chosen.

B



Population	Mean	Median
Single cells, focus & PE+	0.134	0.0939

C



Population	Mean	Median
Single cells, focus & PE+	1.321	1.244

D

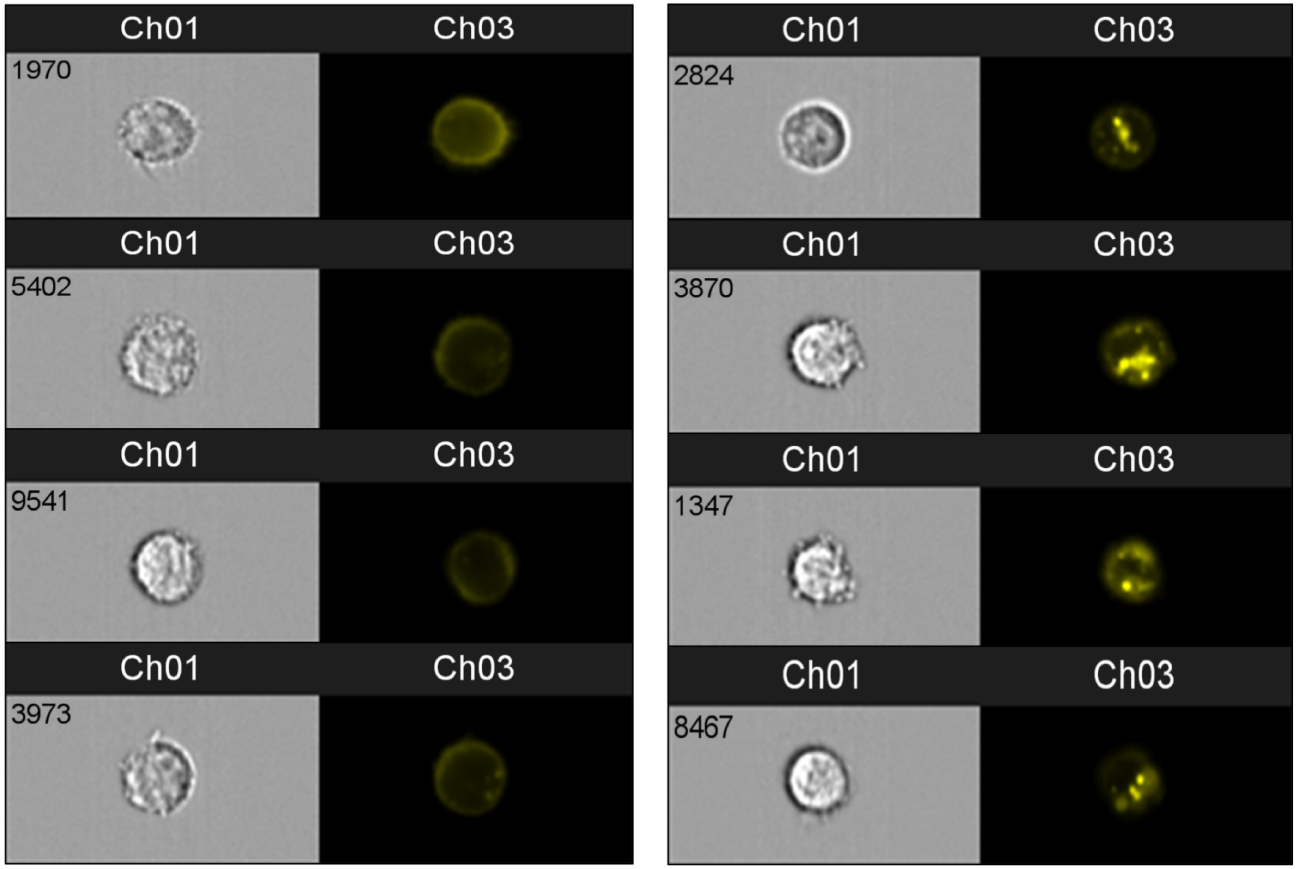


Figure 3.12 Internalisation of αDEC205-PE. A depicts a schematic of the experimental strategy. EBV B cells were pulsed with αDEC205-PE and the degree of internalisation measured using ImageStream X. Single focused and cells positive for PE were gated. Internalisation scores were calculated, reflecting the ratio of cytoplasmic to total brightness intensity. B and C show histogram plots of internalisation scores for cells immediately after pulsing and after incubation for 2 hours at 37°C, respectively. Representative cell images captured by the Amnis ImageStream Flow Cytometer are also provided (D). Selected cell images were chosen to be well above and below the median with the first column showing brightfield images (Ch01) and the second column showing images of fluorescence of αDEC205 (Ch03). Dotted green lines indicated in C correspond to the location on the histogram where representative images were chosen.

3.10 Concluding remarks

In summary, work presented in this chapter demonstrates that conjugation to PWM can be used for delivery of synthetic peptides into an APC for processing and epitope generation. Through a series of experiments, I show that the response to the PWM-based ADS is HLA-restricted and as a result of internalisation of PWM. Using a peptide-specific T cell hybridoma, I also establish that the optimal epitope density present on the cell surface occurs 16 hours after delivery of the ADS. A T cell response was not observed using an αDEC205 delivery system although binding to the cell surface was evident and internalisation can be inferred.

CHAPTER 4: PULSING OF ANTIGEN-PRESENTING CELLS AND PEPTIDE ELUTION TO IDENTIFY NOVEL HYBRID EPITOPES

4.1 Background to the Chapter

The previous chapter established a system based on conjugation to PWM to efficiently deliver a 41-mer peptide containing a known HLA-DR4-restricted epitope into EBV B cells for processing and presentation. This approach benefits from being broadly applicable to any peptide containing a C-terminal cysteine residue, allowing it to be conjugated to PWM. This chapter will exploit this fact to deliver a hybrid polypeptide into EBV B cell lines homozygous for HLA-DR3/DQ2 and -DR4/DQ8. Analysis of the immunopeptidome will elucidate whether hybrid epitopes are generated from the polypeptide species.

Co-elution of the natural BDC-2.5 ligand (WE14) with insulin C-peptide fragments was the prerequisite for identification of the first HIP within mouse β-cell granules, containing the DLQTLAL C-peptide sequence and the WSRM WE14 sequence (Delong et al., 2016). The longest possible species that this epitope could be derived from is the entire C-peptide sequence upstream of DLQTLAL and the WE14 sequence downstream of WSRM (EVEDPQVAQLELGGGPGAGDLQTLAL*WSRMDQLAKELTAE*). The corresponding human sequence: EAEDLQVGQVELGGGPGAGSLQPLAL*WSKMDQLAKELTAE***,** is henceforth termed a C-peptide:WE14 polypeptide in the current study. T cell responses against a hybrid epitope (SLQPLAL*WSKMDQL*) contained within the C-peptide:WE14 polypeptide have been reported in several T1D patients (Baker et al., 2019b). It is tantalising to speculate that this short hybrid epitope derived from APC-dependent processing of the long C-peptide:WE14 polypeptide. The existence of a larger C-peptide:WE14 hybrid species has been further strengthened by identification of a deamidated polypeptide of insulin C-peptide fused to WE14 (EAEDLQVGQVELGGGPGAGSLQPLAL*WSK*) in the human islet proteome

(Peakman & Purcell, unpublished). However, this sequence was cytoplasmic i.e. not HLA-linked and the three amino acids *WSK* could quite possibly originate from many proteins/peptides besides chromogranin A.

Delong and colleagues (2016) also reported reactivity to the DLQTLAL C-peptide region joined to the NAAR sequence from IAPP2. Although the presence of this HIP was confirmed in mouse islet cell extracts, the same antigenic fraction also contained numerous other peptides. Later studies by Wiles et al. (2016) correlated the abundance of this HIP with the magnitude of T cell response (BDC-9.3) providing evidence that this was indeed the antigenic component. T cell responses against the equivalent human sequence (SLQPLAL*NAVEVLK*) are also detected in T1D patients (Baker et al., 2019b). These data induced me to have synthesised a C-peptide:IAPP2 polypeptide (sequence detailed in Materials and Methods) in a similar manner described for the C-peptide:WE14 polypeptide. Preliminary evidence for the existence of a C-peptide:WE14 polypeptide prompted an initial focus on this species for elution studies. For conjugation to PWM via a heterobifunctional crosslinker, a C-terminal cystine residue was added to the C-peptide:WE14 sequence and a N-terminal biotin moiety. The resulting sequence: Biotin-EAEDLQVGQVELGGGPGAGSLQPLAL*WSKMDQLAKELTAE*C, was used for downstream pulsing experiments.

4.2 Pulsing of B cell lines with hybrid polypeptide ADS

EBV B cell lines WT51 and PF04015, homozygous for HLA-DR4/DQ8 and -DR3/DQ2 respectively, were pulsed with PWM conjugated to C-peptide:WE14. A total of 1×10^9 cells/ cell line were pulsed in batches (100-150 $\times$ 10^6 cells/pulse) with an optimised ADS concentration (final peptide concentration 0.1 μM). After each pulsing experiment, an aliquot of cells was taken and stained with an anti-biotin antibody. FACS analysis showed that cells

pulsed with the ADS stained positive for biotin indicated by an increase in biotin MFI compared to the isotype and non-pulsed control (Figure 4.1A). Pulsed cells were incubated for 16 hours (optimal epitope density) at 37°C, 5% CO_2 without washing and then harvested for immunoaffinity purification. Given the plan for downstream purification of HLA-DR and -DQ complexes, after harvesting cells were stained for cell surface expression of HLA-DR and HLA-DQ. Both cell lines co-express high levels of HLA-DR and -DQ with 99.9% of cells being double positive (Figure 4.1B). The data confirms presence of the ADS at the cell surface after pulsing and high expression levels of target HLA molecules to be later purified.

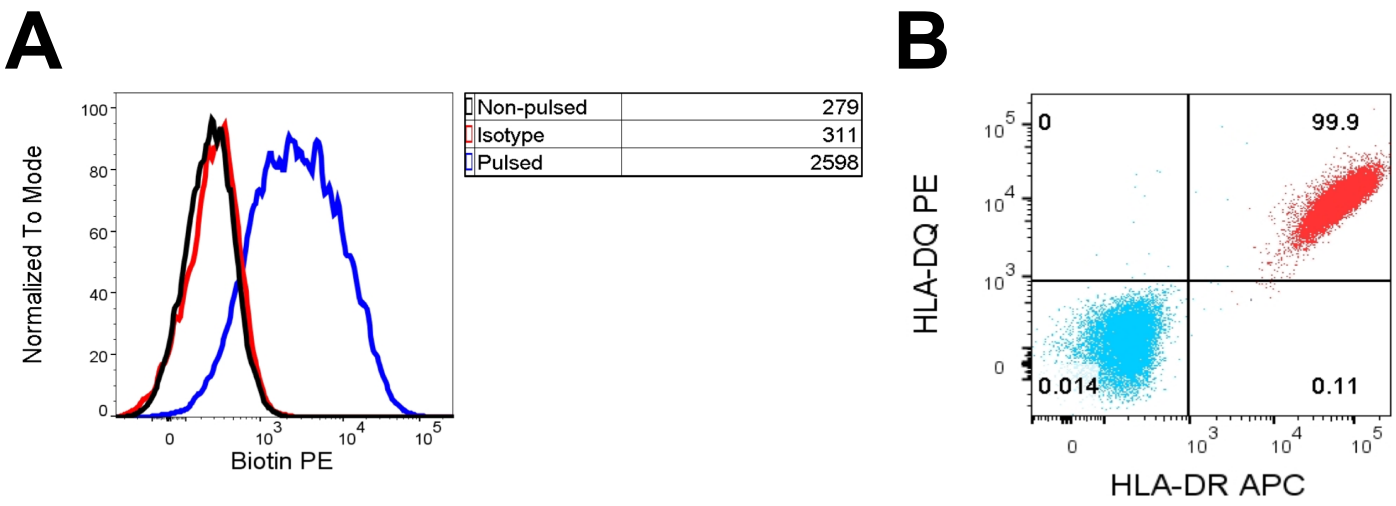


Figure 4.1 Biotin-labelled peptide is present on the cell surface after pulsing and cells express high levels of HLA-DR and -DQ. After pulsing with PWM conjugated to C-peptide:WE14 polypeptide, cells were stained with an anti-biotin antibody (A) and class II HLA antibodies (B). A and B show representative staining from PF04015 and WT51 cells pulsed with the antigen delivery system, respectively. Isotype-control stained cells are shown in both cases (red, histogram; blue, dot plot). Values indicated in the table are the median fluorescence intensities.

4.3 Building class II immunoaffinity columns for elution

To purify HLA class II molecules from pulsed cell pellets, affinity columns containing immobilised anti-DR (clone: L243) and anti-DQ (clone: SPVL3) antibodies were generated. In order to first obtain milligram antibody quantities, L243 and SPVL3 were purified from mouse

ascites. The ascites is conditioned to remove high molecular weight proteins such as transferrin and the conditioned ascites subsequently incubated with a non-IgG binding resin. Purified IgG is collected in the flow through with an identical protein band pattern to commercially available antibody by Coomassie staining (Figure 4.2); with bands corresponding to both heavy (50 kDa) and light chains (25 kDa). Bands at ~28-30 kDa were also observed and most likely represent glycosylation of the light chain (Tachibana et al., 1996).

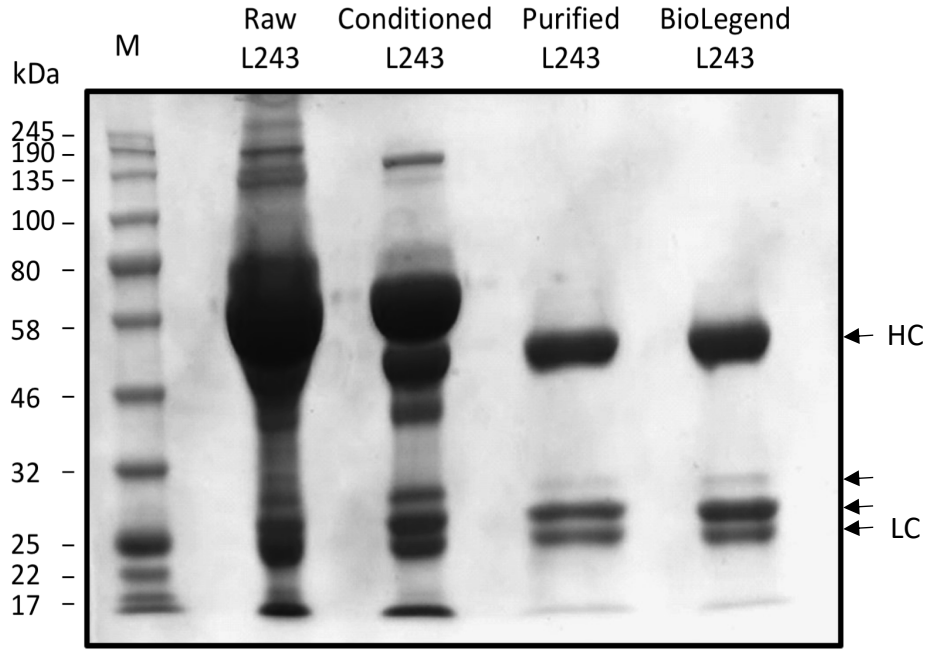


Figure 4.2 Purification of antibody from mouse ascites. Representative Coomassie-stained SDS-PAGE gel (10%) of purification of anti-DR (L243) ascites. Raw ascites was conditioned to remove abundant proteins such as transferrin and then purified using a resin which binds to non-IgG proteins. Purified L243 purchased from BioLegend was also stained for comparison. M, molecular weight standard; kDa, kilodalton; HC, heavy chain; LC, light chain.

Before proceeding to couple purified L243 and SPVL3 antibodies to Protein A Sepharose (PAS) beads, it was first important to eliminate the possibility that the purification process lead to a loss in the activity of the antibody. To test this, EBV B cells were stained with a titration of either purified L243 or SPVL3 antibody. A fluorescently labelled secondary antibody was used to detect binding to the cell surface (Figure 4.3A, C). The same titration was also

performed with commercially available L243 and SPVL3 antibodies to allow comparison of MFI values. Data shows that similar MFI values were observed between purified antibodies and their commercially available counterparts (Figure 4.3B, D), indicating that at a given concentration the amount of anti-DR/DQ antibody binding to the cell surface was similar. These findings illustrate accurate quantification of purified L243 and SPVL3 antibodies and confirm the ability to bind HLA antigen.

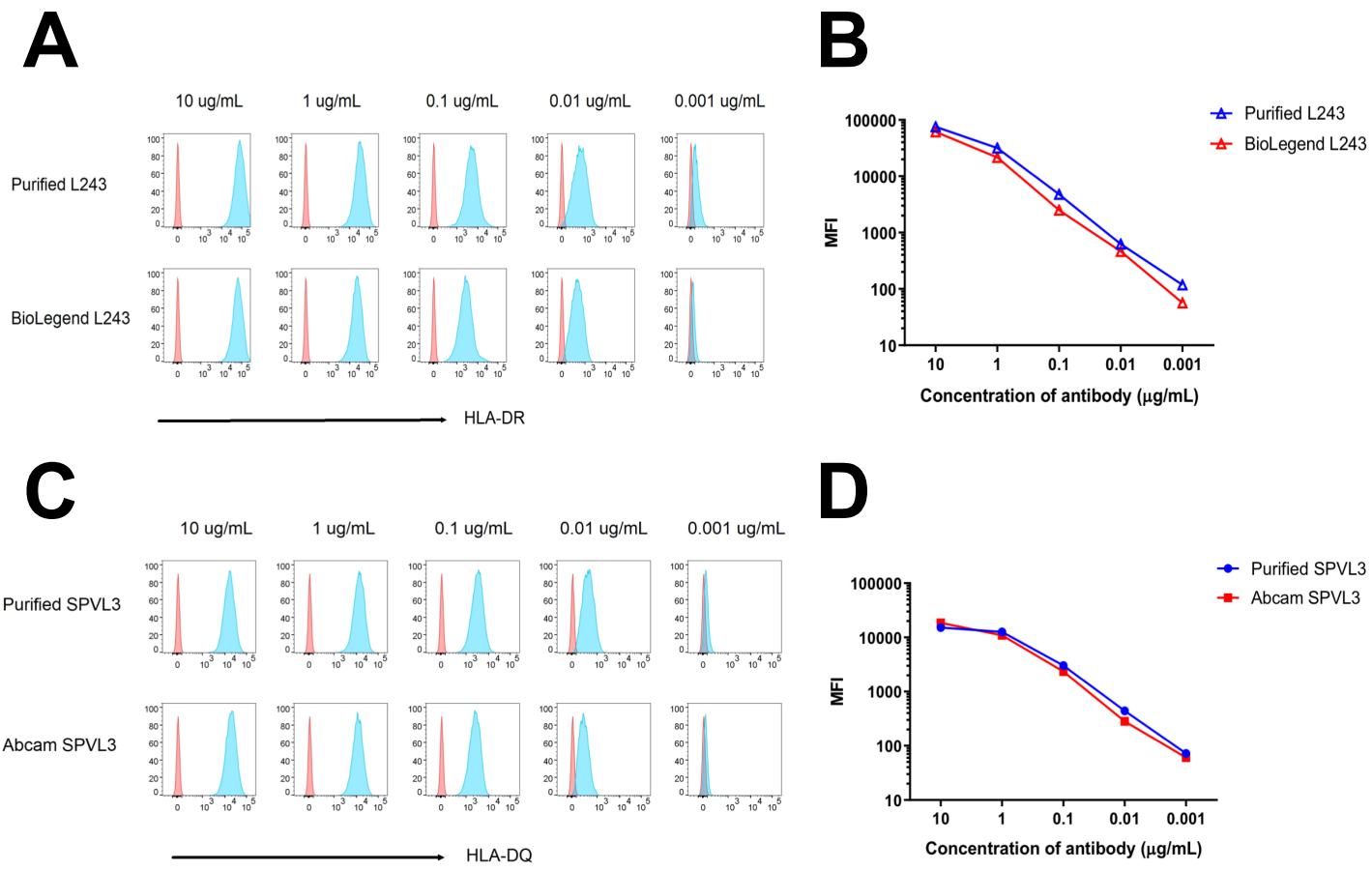


Figure 4.3 Purified antibodies bind HLA. An EBV B cell line was stained with a titration of purified or commercially available anti-DR (L243)/anti-DQ (SPVL3) antibodies. Binding to the cell surface was detected via a fluorescently labelled secondary antibody. A and C display representative histograms for each concentration of antibody relative to an isotype control (red histogram). The corresponding MFI values are shown in B and D. Data are representative for at least two separate experiments.

Whilst PAS beads have high affinity for the Fc region of IgG, dimethyl pimelimidate (DMP) was used to covalently link the Fc region of the antibody heavy chain to Protein A. To monitor the reaction, pre- and post-crosslinked beads were analysed by SDS-PAGE. A small-scale citrate buffer acid elution of the crosslinked beads was also performed. Pre-crosslinking, the antibody is not covalently bound to the beads, thus when the sample is boiled in reducing sample buffer the antibody comes off the beads and the heavy and light chains are visible (Figure 4.4). Successful crosslinking was indicated by the presence of the antibody heavy chain on the affinity bound beads but not on the post-crosslinked beads. The light chain, however, appears in both. Following incubation of antibody with PAS beads, the column flow through was depleted of antibody indicated by the reduced staining intensity compared with the purified antibody loaded on the column. Importantly, citric acid elution did not strip any antibody from the column.

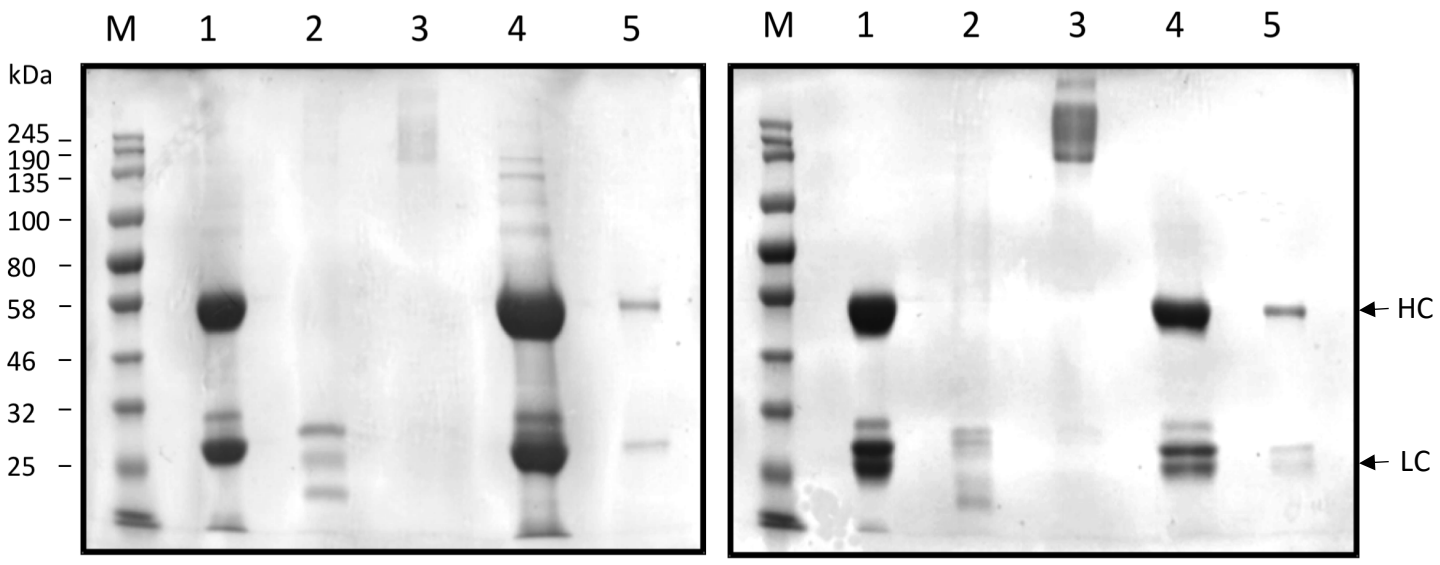


Figure 4.4 Generation of MHC immunoaffinity columns. Representative Coomassie-stained SDS-PAGE gels (4-20%) monitoring crosslinking of SPVL3 (left) and L243 (right) antibodies to Protein A Sepharose beads. Lane 1: Pre-crosslinked beads. Lane 2: Post-crosslinked beads. Lane 3: Citrate buffer elution. Lane 4: Purified antibody. Lane 5: Flow through from bead-antibody incubation. M, molecular weight standard; kDa, kilodalton; HC, heavy chain; LC, light chain.

4.4 Affinity purification of HLA-DR and -DQ complexes

HLA-DR and -DQ molecules were purified from cell pellets pulsed with the ADS as previously described (Purcell et al., 2019a). In brief, cell pellets were homogenised using lysis buffer and sonication. The lysate was pre-cleared by centrifugation to remove the nuclei before loading the supernatant onto a protein A sepharose pre-column lacking crosslinked antibodies (Figure 4.5). The flow through from this pre-column is then loaded directly onto the crosslinked monoclonal antibody columns (L243 and SPVL3) by gravity flow. For maximal yield, the lysate is run through the columns multiple times and the flow through collected at each stage. In western blotting, anti-HLA-DR antibody confirmed binding of class II molecules present in the lysate to the antibody column with each subsequent loading cycle (Figure 4.7A). Peptide-HLA complexes are eluted from the column using 10% acetic acid, causing dissociation from the antibody and also dissociation of the bound peptide.

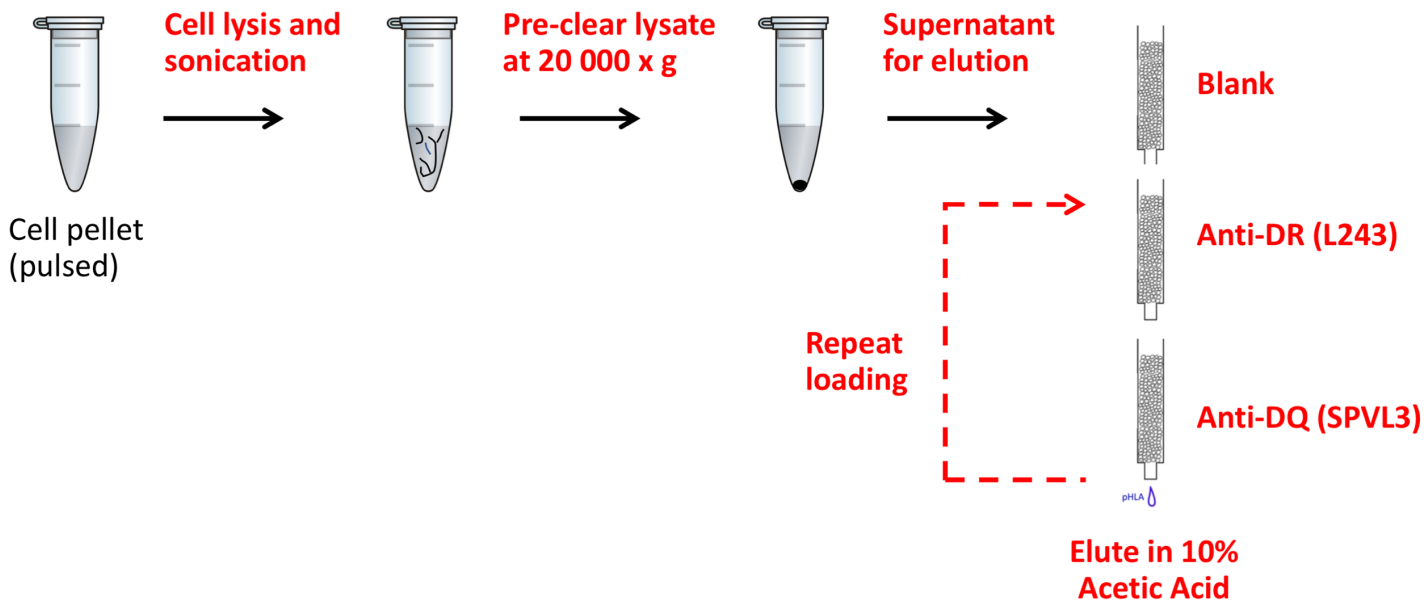


Figure 4.5 Elution of pHLA complexes. The blank column contains protein A sepharose beads (without antibody crosslinking) and is used to remove nonspecific binding proteins from the lysate. All steps are performed under gravity flow. pHLA, peptide-HLA.

4.5 Isolation of bound peptides and mass spectrometry

To enrich peptides (separate them from the HLA class II molecules) the acid eluate was loaded onto a C18 column containing a hydrophobic surface of silica particles (Figure 4.6). Bound peptides were eluted using 30% acetonitrile (ACN) containing 0.1% trifluoroacetic acid (TFA) and analysed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). As a quality control step the remaining proteins bound to the C18 column, which should include HLA, were eluted in 80% ACN:0.1% TFA. Western blot analysis revealed the presence of HLA-DR and -DQ protein in the 80% ACN eluate (Figure 4.7B). The flow through after loading of the acetic acid eluate onto the C18 column was also probed for HLA-DR and -DQ protein. DR but not DQ protein was detected in the flow through; this was, however, marginal compared to the amount of DR protein eluted from the C18 column with 80% ACN suggesting that the majority of DR protein did indeed bind to the column.

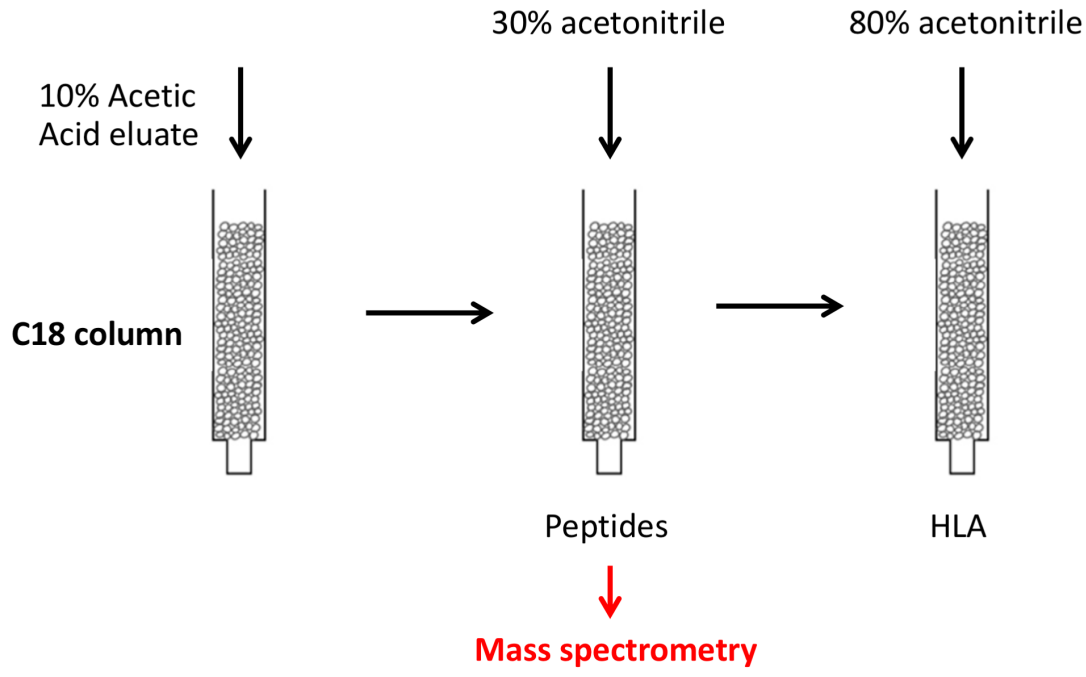


Figure 4.6 Peptide clean-up of MHC eluate by C18 column. Peptides eluted from the column were vacuum-concentrated to dryness before being re-constituted with 2% acetonitrile (ACN)-0.1% formic acid (FA) prior to mass spectrometric analysis. HLA, human leukocyte antigen (column image courtesy of Deborah Kronenberg).

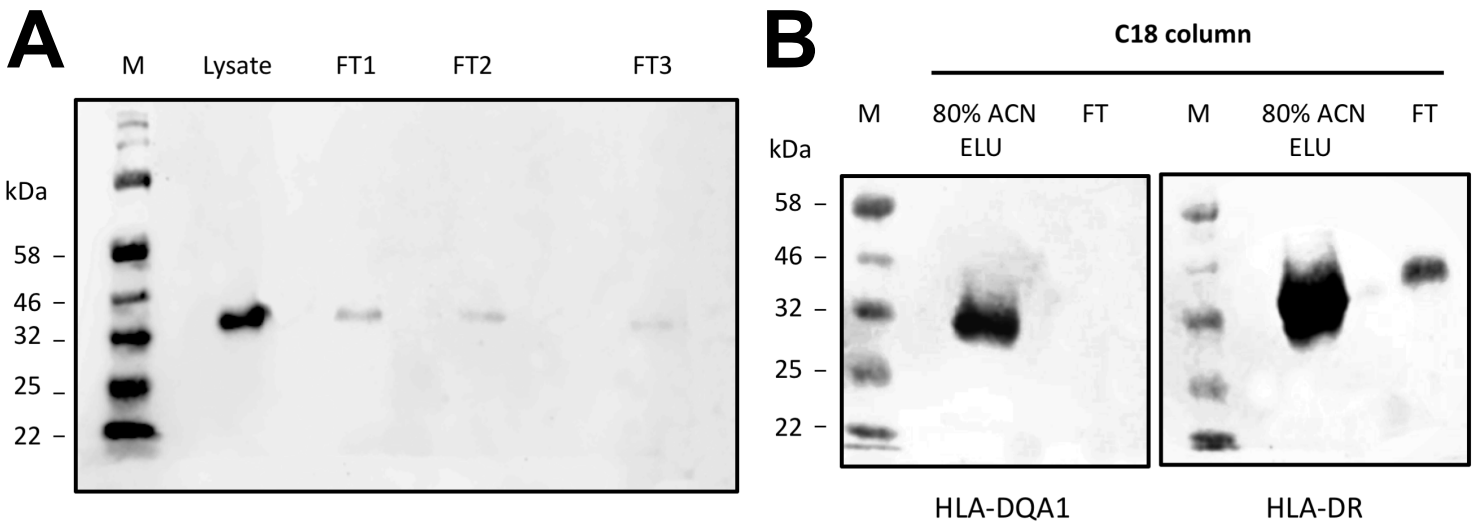


Figure 4.7 Depletion of flow through and presence of HLA-II in 80% acetonitrile eluate. Western blot analysis of HLA-DR protein in pulsed WT51 cell lysate and after each successive loading cycle (A). B shows analysis of flow through and 80% acetonitrile eluate from C18 column using antibodies to detect HLA-DQA1 (left) and HLA-DR (right). Blots shown in B represent material obtained by peptide elution of pulsed PF04015 cells. FT, flow through; M, molecular weight standard; kDa, kilodalton; ACN, acetonitrile; ELU, eluate.

HLA class II complexes were purified from whole cell lysate using 7×10^8 pulsed cells as the starting material. This reserved 3×10^8 cells for validation of targets should putative epitopes be identified. MS analysis was performed using two approaches: PD (ProteomDiscoverer 2.2) (conventional) and PEAKStudio 7.5 (Database Search and *de novo* sequencing). Peptides were first matched against the Uniprot Swiss-Prot human proteome database plus peptide sequences of interest, i.e. the C-peptide:WE14 polypeptide used for pulsing and PWM to identify any pokeweed-derived peptides. To accurately identify peptides, a false discovery rate (FDR) threshold of less than 1% was applied. The total number of identified peptides from HLA-DR4 was 6376, with similar data for -DQ8 (Table 4.1). In comparison, fewer peptides were identified from -DR3 and -DQ2 with the least number of peptides from -DQ2 identified overall (3078 peptides). Thus, more than 20 000 unique peptides were identified across the four analysed HLA molecules.

Table 4.1 Number of unique peptides identified for each HLA allele with a threshold of peptide FDR less than 1%.

	FDR < 1%
*HLA-DRB1*04:01* (DR4)	6376
*HLA-DQA1*03:01; DQB1*03:02* (DQ8)	6256
*HLA-DRB1*03:01* (DR3)	5212
*HLA-DQA1*05:01; DQB1*02:01* (DQ2)	3078
All HLA	20 922

†HLA, human leukocyte antigen; FDR, false discovery rate.

4.6 Length of eluted peptides and identification of nested sets

Initial scanning of the complete dataset did not identify novel hybrid epitopes. To evaluate whether the elution was successful, downstream analysis was performed on a filtered dataset. Analysis was limited to peptides present at a peptide length frequency greater than 1%: at least eight (HLA-DR3 and -DQ2) or nine (HLA-DR4 and -DQ8) residues. Peptides not being part of nested sets were included as many non-nested peptides are predicted to bind their cognate class II molecule.

The peptide length distribution of eluted peptides satisfying these criteria was characteristic of class II peptides (Lippolis et al., 2002), varying between 8 and 24 residues with a peak in distribution at 15 amino acids (Figure 4.8). For all HLA molecules, the next most common peptide length was 16 amino acids. For HLA-DQ8, -DR3 and -DQ2 14-mer long peptides represented the third most abundant length whereas for HLA-DR4, this was 17-mer peptides. An interesting peak of 9-mers was present in the HLA-DQ2 elution; closer analysis showed that these peptides lack a -DQ2 binding motif and did not form nested sets. This peak at 9 amino acids was, however, less pronounced for the other alleles suggesting that this may be associated

with fewer peptides being eluted from -DQ2 molecules. This finding is not surprising given identification of the unusually short PPI_{17-24} (8-mer) epitope from heterozygous HLA-DQ2/8 expressing DCs (van Lummel et al., 2016b). This PPI epitope also does not fulfil the high-risk HLA-DQ binding motif. Most of the eluted peptides could be assigned to nested sets of sequences (Figure 4.9A, B, Table 4.2 and 4.3) in agreement with previous reports (Chicz et al., 1993; Peakman et al., 1999; van Lummel et al., 2016b; Zhou et al., 2017b). Nested sets are groups of peptides with a consensus core binding motif but varying extended N- and/or C-termini and are typical of processing through the class II pathway (Chicz et al., 1992).

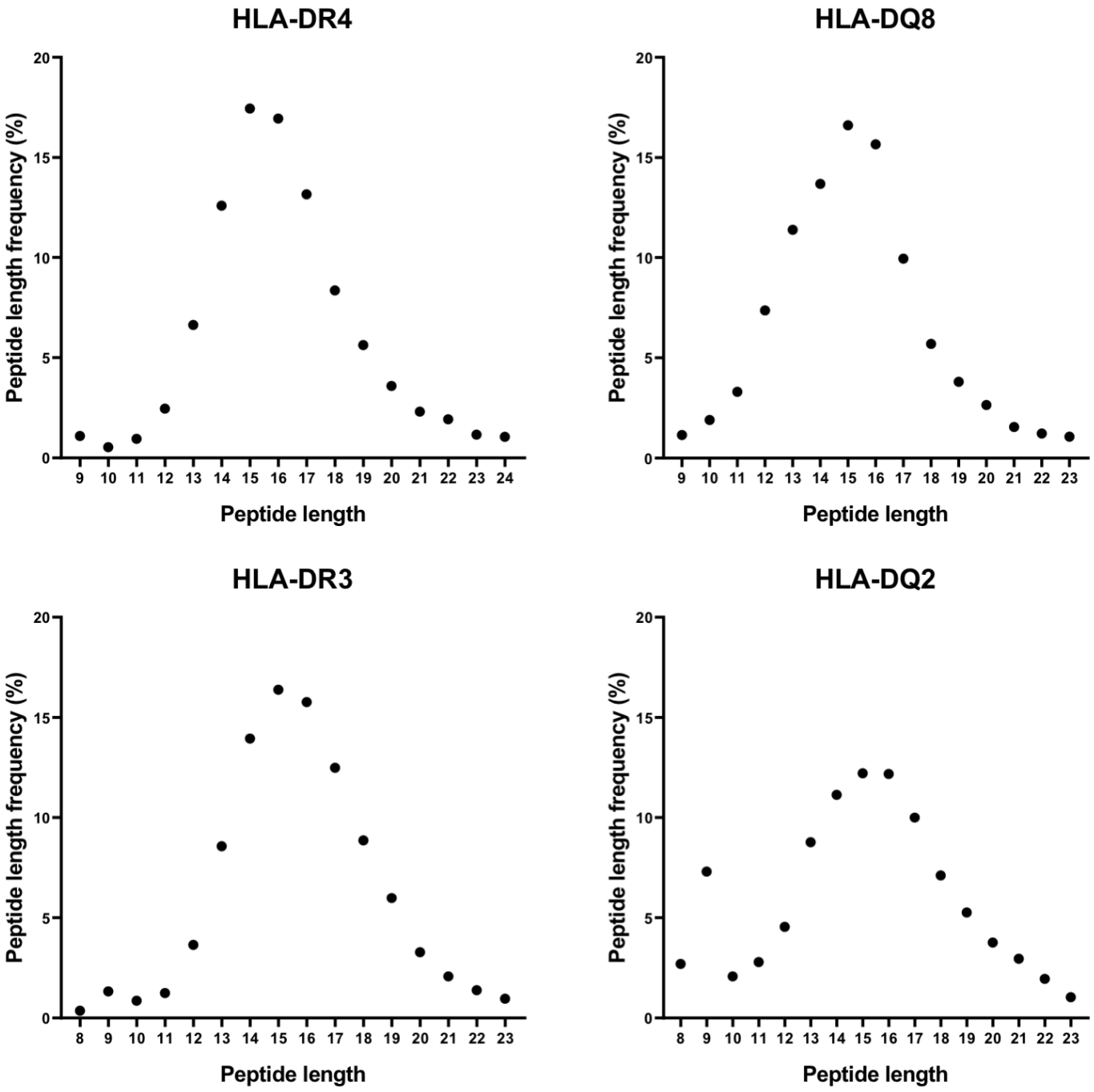


Figure 4.8 Peptide length distribution. The peptide length distribution of peptides identified with a peptide FDR <1% is shown for peptides with a frequency greater than 1%. Each subsequent increment in length is shown until the frequency falls below 1%. Predominant peptide length is 14–17-mer with 15-mer being the most abundant length for all analysed HLA molecules. HLA, human leukocyte antigen; FDR, false discovery rate.

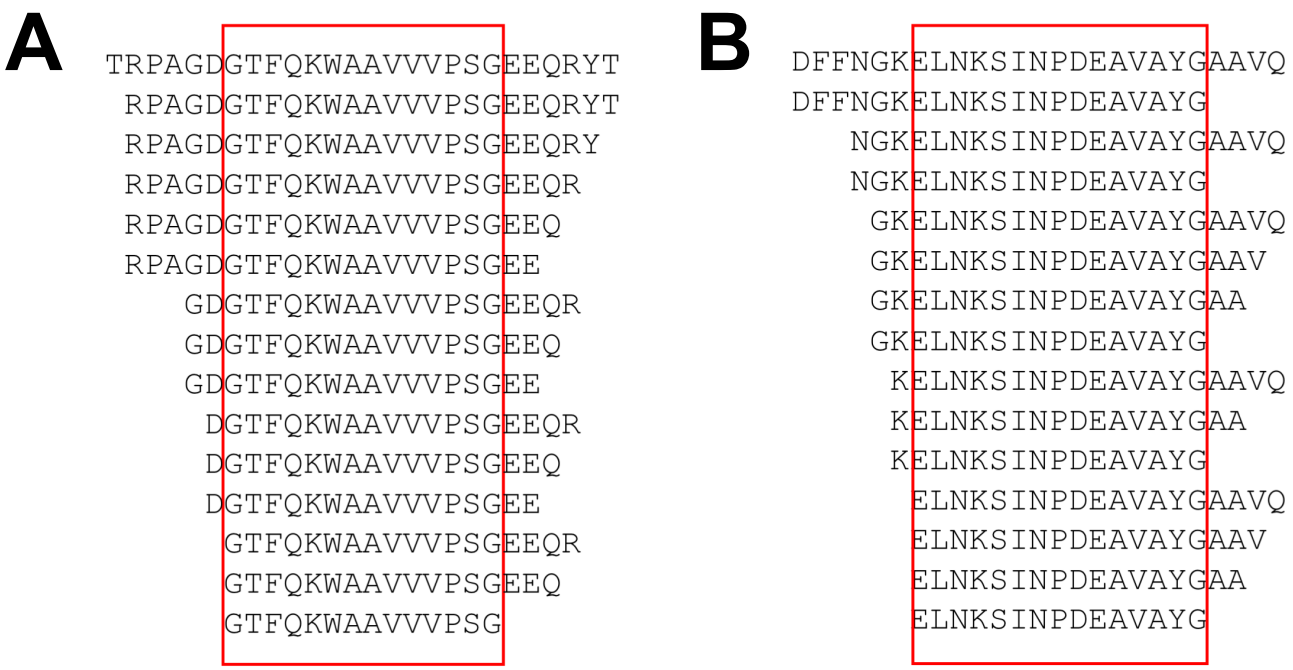


Figure 4.9 Peptides are eluted in nested sets. Representative nested set peptides eluted from HLA-DQ8 (A) and -DR3 (B). Peptides derived from HLA class I and heat shock protein are shown by A and B, respectively. Core sequences are indicated with red boxes.

Table 4.2 Frequency of nested peptide sets eluted from HLA-DR4.

Number of peptides from source protein	Number of proteins identified by X peptides	Frequency of proteins identified with nested peptide sets (%)	Frequency of proteins identified without nested peptide sets (%)
2	205	76.1	23.9
3	122	95.1	4.9
4	86	100	0
5	59	98.3	1.7
6	40	100	0
7	34	100	0
8	26	100	0

†Nested sets were defined by the presence of at least two peptides which have a nested sequence. To identify a nested set, the following criterion was applied: the sequence of the shorter peptide is 100% represented in the longer peptide and the shorter peptide is at least nine amino acids in length. Data shown in the table is limited to the number of peptides represented by greater than 20 proteins. Similar results were obtained for peptides eluted from HLA-DR3.

Table 4.3 Frequency of nested peptide sets eluted from HLA-DQ8.

Number of peptides from source protein	Number of proteins identified by X peptides	Frequency of proteins identified with nested peptide sets (%)	Frequency of proteins identified without nested peptide sets (%)
2	177	72.9	27.1
3	93	88.2	11.8
4	67	89.6	10.4
5	44	100	0
6	50	98	2
7	25	100	0
8	35	100	0

†Nested sets were defined by the presence of at least two peptides which have a nested sequence. To identify a nested set, the following criterion was applied: the sequence of the shorter peptide is 100% represented in the longer peptide and the shorter peptide is at least nine amino acids in length. Data shown in the table is limited to the number of peptides represented by greater than 20 proteins. Similar results were obtained for peptides eluted from HLA-DQ2.

4.7 Binding motif

Determination of the 9-mer core region of the most abundant peptide group (15-mer), reported to be essential for HLA-II binding (Madden, 1995; Rammensee et al., 1995), revealed allele-specific amino acid preferences (Figure 4.10). In accordance with previous literature, four anchor residues were observed for binding HLA class II molecules locating to positions (P) 1, 4, 6 and 9 (Sant'Angelo et al., 2002). For HLA-DR4, I identified a preference for a bulky aromatic (phenylalanine, tyrosine, tryptophan) or an aliphatic residue (leucine, isoleucine) at P1 and a preference for a negatively charged amino acid (aspartate and glutamate) at P4. P6 of core 9-mer -DR4 sequences displayed a preference for polar residues (threonine, serine and asparagine) while P9 showed an equal preference for polar (serine, glutamine) and hydrophobic residues (alanine, glycine). Similarly, the HLA-DR3 binding motif featured a preference for aliphatic residues (isoleucine, leucine and valine) at P1 and a strong preference for aspartate at P4. This is consistent with the observation that epitopes with a negatively charged residue in

position 4 of the binding motif bind more strongly to HLA-DR3 (Inaba et al., 2010). However, unlike HLA-DR4 9-mer core sequences, HLA-DR3 binding cores exhibited a basic residue (lysine or arginine) at P6 and P9.

In line with previous literature on the -DQ8 binding motif (Moustakas et al., 2000), HLA-DQ8 showed a distinct binding preference for negatively charged amino acids (aspartate and glutamate) at P1 and P9 and a non-bulky hydrophobic amino acid at P4 and P6 (alanine or valine). Strikingly, the -DQ2 binding motif was less clear although a strong preference for aspartate and particularly glutamate was found not only at the P6 and P9 anchor positions, but also at many positions in the binding core including the non-anchor positions P2, P5, P7 and P8, consistent with reported findings (Stepniak et al., 2008). The same analysis was also performed for the three most abundant peptide lengths and overall the binding motif appears conserved through the peptide lengths.

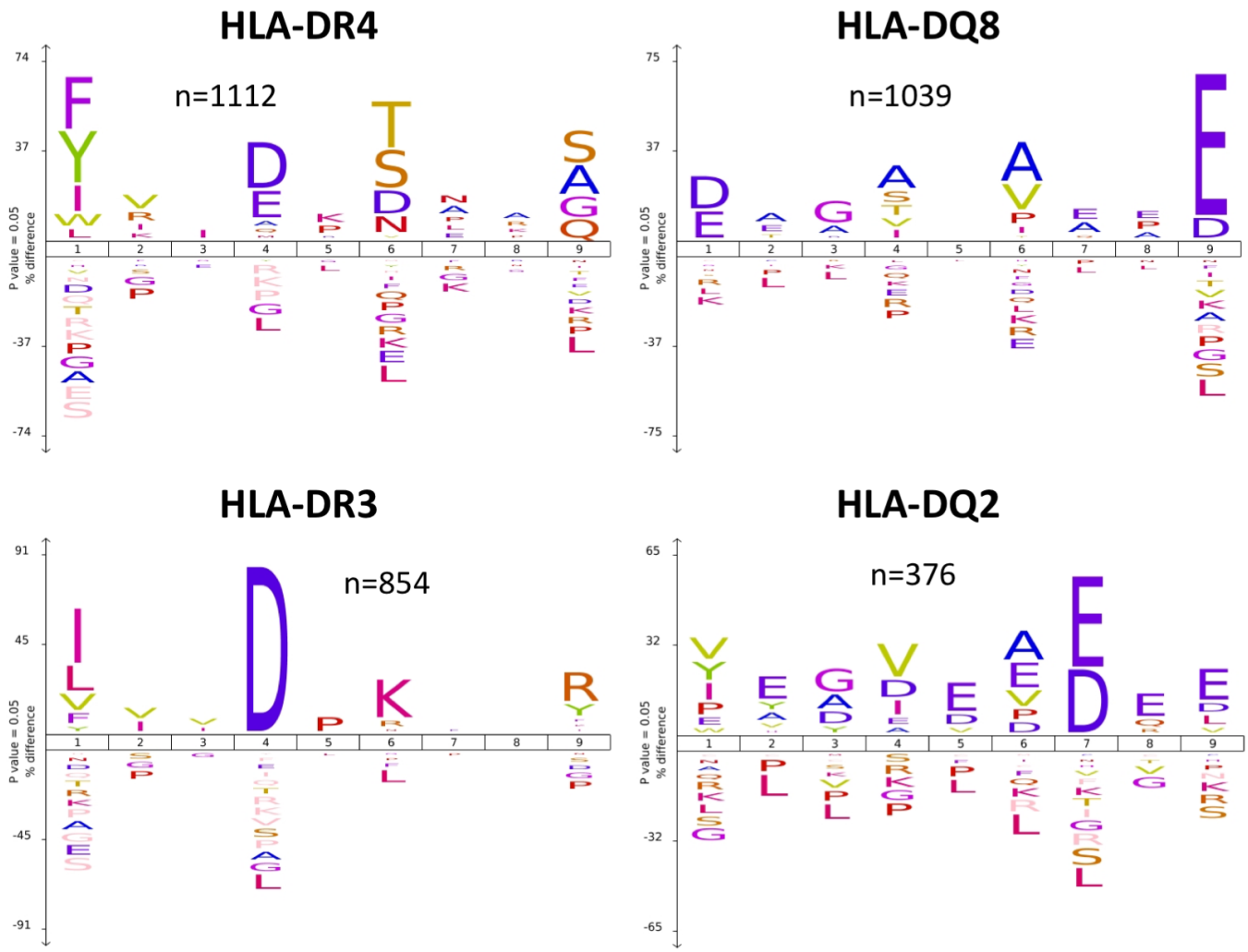


Figure 4.10 Peptide binding motif of core sequences. Sequence Logo plots are shown as examples for the most frequent peptide group (15-mer) of each HLA allele with number of detected 15-mer peptides indicated. The height of each letter within the column is proportional to its frequency at the given position. Predicted binding cores were identified using the NetMHCII 2.3 program and motifs visualised using iceLogo. HLA, human leucocyte antigen.

4.8 Source protein of presented peptides

To examine whether specific source proteins feature strongly in the elution dataset, the source proteins of the twenty most abundant peptides identified from each allele were identified. A list of source proteins identified from peptides eluted from HLA-DR4 is shown in Table 4.4. Similar results were obtained for HLA-DQ8, -DR3 and -DQ2 with CD74 (CLIP), immunoglobin, β_2-microglobin, HLA class I and class II (DQA, DRA, and DRB) peptides being amongst the top twenty most abundant. The most abundant peptide eluted from HLA-

DR3 was a peptide derived from the transferrin receptor. These findings are in agreement with previously reported source protein data of class II peptides (Bergseng et al., 2015; Wahlstrom et al., 2007).

Table 4.4 Twenty most abundant peptides eluted from HLA-DR4.

Peptide	#PSMs	Source protein
GPTTYKVTSTLTIKE	293	Immunoglobulin heavy constant mu
NPRKFNLDATELSIRK	127	Syntaxin-6
DLRSWTAVDTAAQISEQ	111	HLA class I histocompatibility antigen, alpha chain E
KVNLLKIKTELCKKEV	102	Golgi apparatus protein 1
GPTTYKVTSTLTIKES	102	Immunoglobulin heavy constant mu
KVNLLKIKTELCKKE	98	Golgi apparatus protein 1
KPPKPVSKMRMATPLLMQALP	92	HLA class II histocompatibility antigen gamma chain (CD74)
SHRSYSCQVTHEGSTVE	87	Immunoglobulin lambda-1 light chain
SHRSYSCQVTHEGSTVEK	75	Immunoglobulin lambda-1 light chain
YDHNFVKAINAIQKSW	66	Dipeptidyl peptidase 1
QRKWEAARVAEQLRAY	66	HLA class I histocompatibility antigen, B alpha chain
TQRKWEAARVAEQLRAY	62	HLA class I histocompatibility antigen, B alpha chain
KVNLLKIKTELCKKEVL	61	Golgi apparatus protein 1
HPINEYYIADASEDQVFV	58	Sortilin-related receptor
SHRSYSCQVTHEGST	56	Immunoglobulin lambda-1 light chain
LPKPPKPVSKMRMATPLLMQALP	54	HLA class II histocompatibility antigen gamma chain (CD74)
RASWRIISSIEQKEE	52	14-3-3 protein epsilon
HPINEYYIADASEDQVFVC	51	Sortilin-related receptor
YDHNFVKAINAIQKSWT	49	Dipeptidyl peptidase 1
VDDTQFVRFDSDAASQRMEP	48	HLA class I histocompatibility antigen, A alpha chain

[†]High abundance peptides were identified by the number of peptide-to-spectrum matching (PSM) events. The source protein of each peptide is indicated.

4.9 Overlap in the immunopeptidome between different elution studies

To provide even stronger evidence that the inability to identify a hybrid epitope was not due to a technical failure within the elution, I sought to investigate whether peptides eluted in the current study have also been identified in pre-existing datasets. Bergseng et al. (2015) generated a single database containing HLA-DQ2-bound peptides eluted from three different EBV B cell

lines, including PF04015. As the choice of cell line impacts the peptide repertoire (Abelin et al., 2019), the use of a cell line identical to that used in the current study prompted an assessment of overlap between the two datasets (Figure 4.11A). In this case, overlap was defined as identification of a peptide with an identical sequence in both databases. Of the 3078 unique peptides identified in the current study, 1128 (36%) were also identified in the database generated by Bergseng et al. (2015). Considering that the database created by Bergseng and colleagues is larger (4266 unique peptides) and contains DQ2-eluted peptides not only from PF04015 but also two other HLA-DQ2-positive EBV B cell lines, this level of overlap is justifiable. Overall, these data indicate that a proportion of peptides eluted from HLA-DQ2 in the current study have been previously identified.

To bolster this finding, the cellular localisation of the peptide source proteins eluted from HLA-DR4 molecules in the current study was compared to data obtained from DCs homozygous for HLA-DR4 (Figure 4.11B). Comparing the subcellular location to that of the human proteome, epitopes from lysosomal, Golgi apparatus, endosomal membrane, endoplasmic reticulum and cytoplasmic proteins were enriched for both datasets. Similarly, a lower proportion of nuclear and mitochondrial proteins was observed. Interestingly, in DCs, epitopes were enriched for secreted source proteins (27% vs 12% in the human proteome), but this was absent in EBV B cells (14%). The data point toward enrichment of proteins derived from similar source locations, which in combination with an overlap of identified epitopes, undeniably demonstrates successful purification of HLA class II complexes.

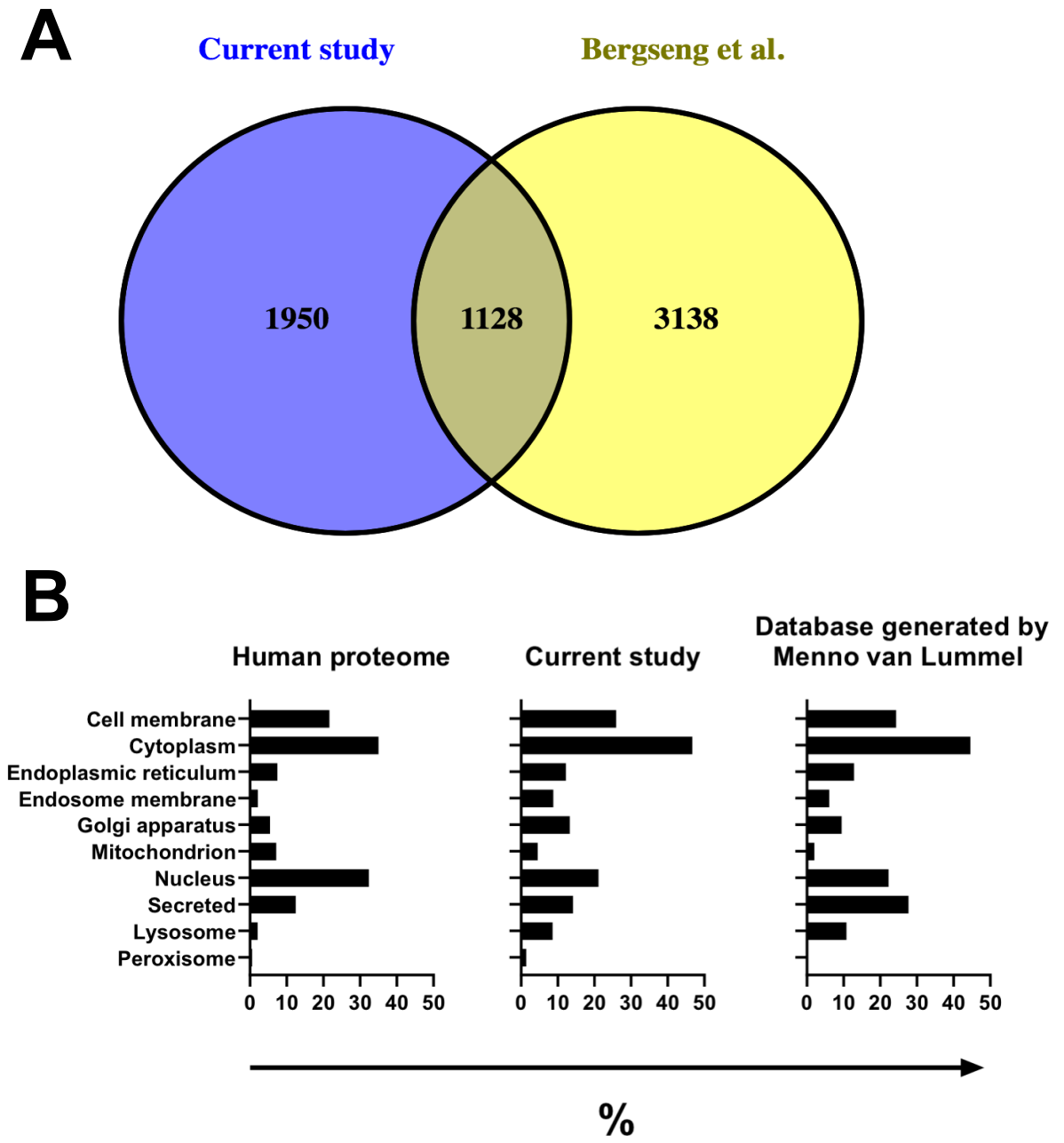


Figure 4.11 Overlap of identified epitopes and source protein location with pre-existing datasets. Venn diagram showing overlap of peptides eluted from HLA-DQ2 molecules between the current study and Bergseng et al. (2015) (A). The number of unique peptides and those peptides common between the two databases are indicated. Dataset created by Bergseng et al. was derived from three cell lines expressing -DQ2.5: CD114 (obtained from a coeliac disease patient), STEINLIN and PF04015. Databases were searched for only identical peptide sequences. B shows the subcellular location of source proteins of peptides eluted from HLA-DR4 molecules in the current study (EBV B cells) and dendritic cells homozygous for HLA-DR4 (database courtesy of Menno van Lummel, Leiden University Medical Centre, unpublished). A graph of this for the human proteome is also shown. Data expressed as percentage of total identified proteins with subcellular location annotations (multiple protein subcellular localisation sources are possible).

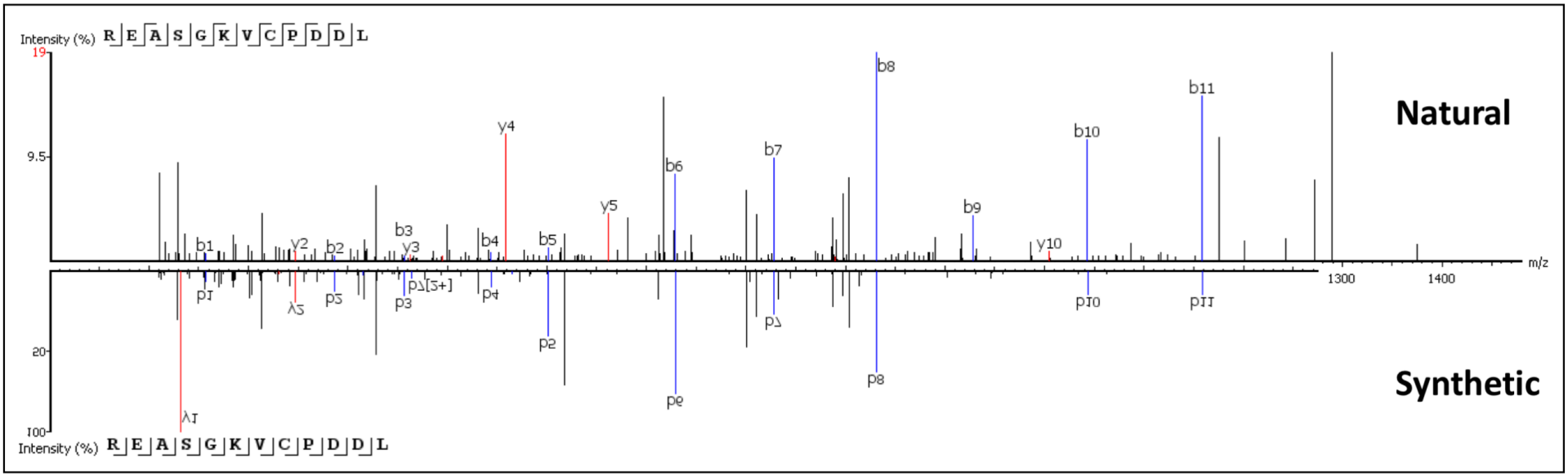
4.10 Identification of epitopes derived from exogenously acquired proteins

To reject the possibility that hybrid epitopes were not identified because of an incapacity of the cell line to process and present exogenous antigen under the pulsing conditions, I searched the mass spectrometry data for exogenously derived peptides. While a hybrid epitope was not identified, it is possible to envisage that epitopes derived from pokeweed mitogen may be generated. Crude PWM is heterogenous with multiple isoforms present (Yamaguchi et al., 2004), any of which could be found within the elution sample. Protein sequences of Lectin-C, -D2, -B and -A are available in the UniProt database and were used for database searching. Identification of a peptide eluted from HLA-DQ8 (REASGKVCPDDL) that derived from Lectin-B provided reassurance at a technical level that the ADS was successful in delivery of the conjugate into the endocytic pathway for epitope generation.

To confirm the identity of the natural pokeweed peptide in the sample, the synthetic peptide REASGKVCPDDL was analysed by MS/MS. Although the MS/MS fragment spectra varies in terms of the fragment ion intensities (Figure 4.12A), a spectrum is validated based on mapping of fragment ions between synthetic and natural samples. This is mainly because many factors might affect the intensity such as different types of instruments and matrix effects (Zhou et al., 2017a), although the same fragmentation method was used in the analysis of both peptides. Figure 4.12B shows that the majority of fragment ions are overlapped with all the b ions being identified either as singly or doubly charged ions. Also important to consider is the retention time of both peptides, i.e. when they are eluted from the column. Both the natural and synthetic pokeweed peptide had a similar retention time of ~20 minutes. Taken together, the synthetic peptide is able to validate the same peptide present in the sample. Additionally, as cells were cultured in medium containing 10% FCS the main component of which is BSA, the elution dataset was also queried for peptides derived from BSA. Among the peptides eluted from HLA-

DQ8, -DR4 and -DR3 molecules, overlapping peptides derived from BSA were indeed present. An example of a BSA peptide eluted from HLA-DR3 is shown in Figure 4.12C, where >85% of the b and y ions were identified (Figure 4.12D). Altogether, the data provide strong evidence of processing and presentation of exogenous antigens.

A



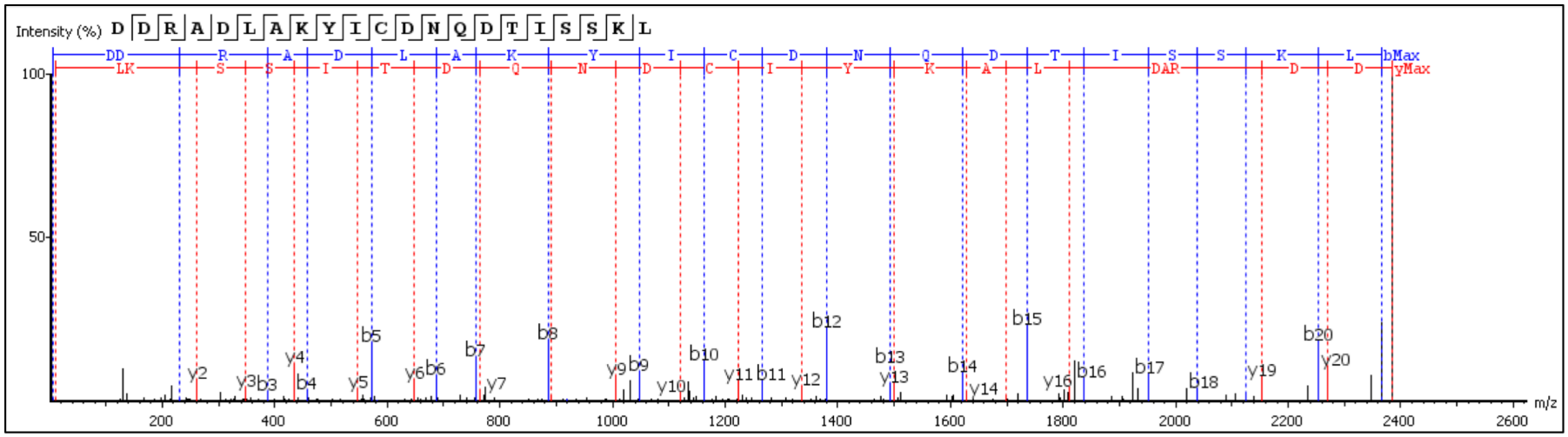
B

	Natural					
#	b^+	b^{2+}	Seq	y^+	y^{2+}	#
1	157.10	79.05	R			12
2	286.15	143.58	E	1133.51	567.26	11
3	357.19	179.09	A	1004.48	502.74	10
4	444.22	222.61	S	933.43	467.22	9
5	501.24	251.12	G	846.40	423.70	8
6	629.34	315.17	K	789.38	395.19	7
7	728.41	364.70	V	661.29	331.14	6
8	831.41	416.21	C	562.22	281.61	5
9	928.46	464.73	P	459.21	230.10	4
10	1043.50	522.25	D	362.16	181.58	3
11	1158.52	579.76	D	247.13	124.06	2
12			L	132.10	66.55	1

	Synthetic					
#	b^+	b^{2+}	Seq	y^+	y^{2+}	#
1	157.11	79.05	R			12
2	286.15	143.58	E	1133.51	567.26	11
3	357.19	179.09	A	1004.47	502.74	10
4	444.22	222.61	S	933.43	467.22	9
5	501.24	251.12	G	846.40	423.70	8
6	629.34	315.17	K	789.38	395.19	7
7	728.40	364.71	V	661.29	331.14	6
8	831.41	416.21	C	562.22	281.61	5
9	928.47	464.73	P	459.21	230.12	4
10	1043.49	522.25	D	362.16	181.58	3
11	1158.52	579.76	D	247.13	124.06	2
12			L	132.10	66.55	1

Figure 4.12 Exogenous proteins are processed and presented. A shows an inverse spectrum comparison of the MS/MS product ion spectra for REASGKVCPDDL peptide identified by elution of HLA-DQ8 (natural) and the synthetic version. Mass-to-charge (m/z) values for fragment ions of types b (blue) and y (red) are given for the natural and synthetic peptide in the ion table (B). Singly and doubly charged b and y ions are indicated. Where the observed and predicted m/z values match, the fragment ions are coloured.

C



D

#	b	Seq	y	#
1	116.03	D		21
2	231.06	D	2269.11	20
3	387.16	R	2154.08	19
4	458.20	A	1997.98	18
5	573.23	D	1926.95	17
6	686.31	L	1811.92	16
7	757.35	A	1698.83	15
8	885.44	K	1627.80	14
9	1048.50	Y	1499.70	13
10	1161.59	I	1336.64	12
11	1264.60	C	1223.56	11
12	1379.63	D	1120.55	10
13	1493.67	N	1005.52	9
14	1621.73	Q	891.48	8
15	1736.75	D	763.42	7
16	1837.79	T	648.39	6
17	1950.88	I	547.34	5
18	2037.92	S	434.26	4
19	2124.93	S	347.23	3
20	2253.04	K	260.20	2
21		L	132.10	1

Figure 4.12 Exogenous proteins are processed and presented. A representative MS/MS spectra and corresponding fragment ion table of a bovine serum albumin peptide identified by elution of HLA-DR3 molecules is shown by C and D, respectively. Only singly charged b and y ions are shown in the ion table. Where the observed and predicted m/z values match, the fragment ions are coloured.

4.11 Possible reasons for the failure to identify a hybrid epitope

4.11.1 The hybrid epitope is a non-binder

One possible explanation for not eluting a hybrid epitope from polypeptide-pulsed BLCLs is that the epitope does not bind to T1D-associated class II HLA alleles (HLA-DR4, -DR3, -DQ8 and -DQ2). To investigate this, the binding affinity of overlapping 15-mer peptides encompassing the C-peptide:WE14 polypeptide (termed peptides 1 to 27) were predicted using the NetMHCII 2.3 algorithm. The algorithm computes a percentile rank (% rank) for a peptide, generated by comparing its score against the scores of a set of 1×10^6 random natural peptides of the same length of the query peptide (Jensen et al., 2018). Because the peptide–MHC affinity may vary widely for different MHC molecules, % rank is considered a better measure to define binding peptides. MHC II peptides with a rank <2 are considered strong binders and those with a % rank <10 are considered weak binders, as described previously (Mason et al., 2019). A fragment of IA-2 comprising residues 709-732 has been previously identified as a naturally processed epitope shown to bind strongly to HLA-DR4 (Peakman et al., 1999); reassuringly this peptide was predicted to have high affinity for HLA-DR4 using the algorithm (% rank = 0.3).

Figure 4.13 indicates the % rank of each 15-mer peptide plotted for each of the T1D-risk class II alleles. Data shows the majority of overlapping peptides, including those where the core sequence contains the hybrid junction, have a predicted % rank greater than 10% suggesting that these peptides do not bind. However, overlapping peptides 1 and 2 are predicted to bind HLA-DQ8 with high affinity (% rank <1) while peptide 3 binds with weak affinity (% rank = 2.5). Importantly, the amino acid sequence of peptides 1-3 contain only the C-peptide region, devoid of any hybrid component. In general, most overlapping peptides are not predicted to

bind with high affinity to the four analysed HLA molecules, supporting the notion that a hybrid epitope was not eluted due to a lack of binding capacity.

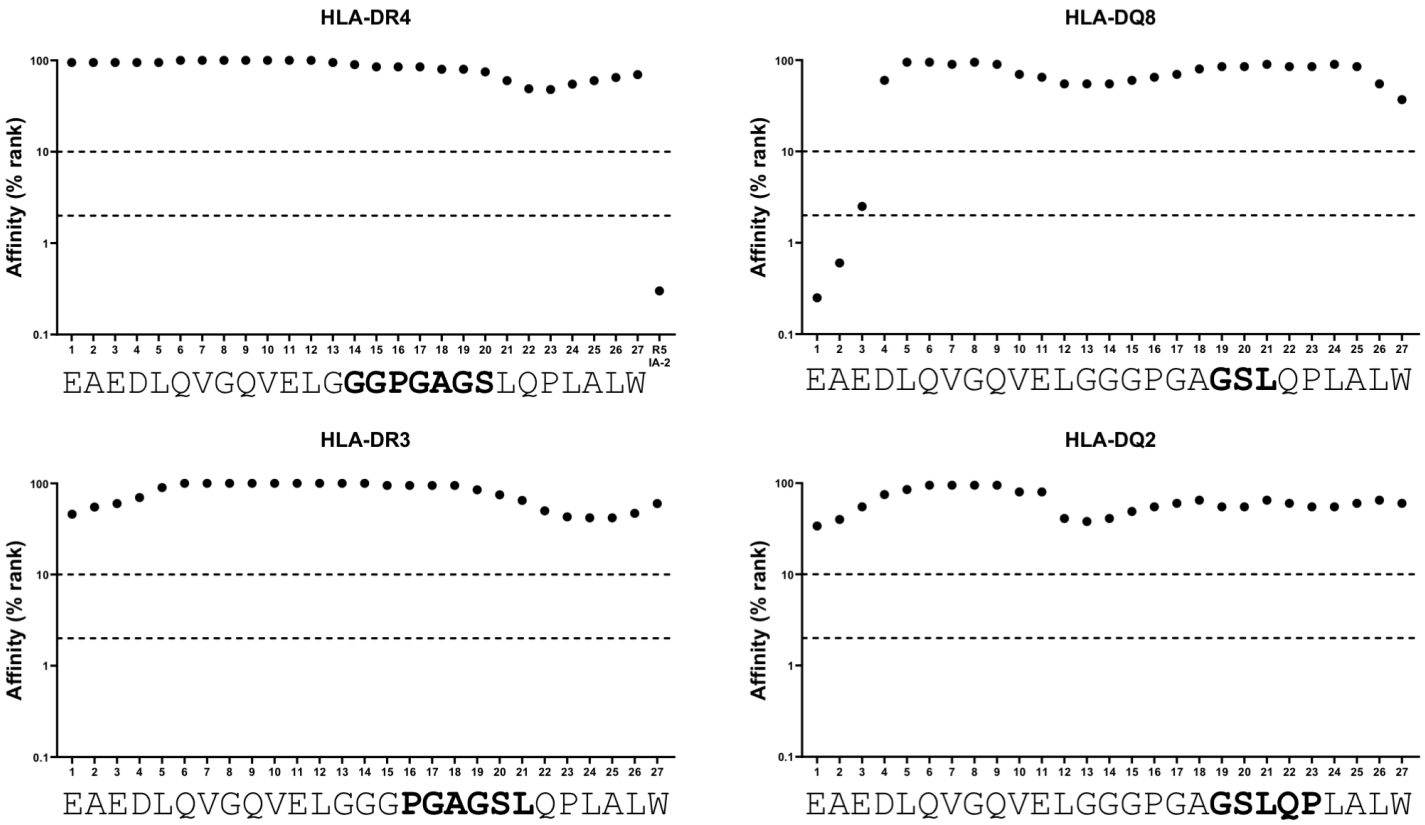


Figure 4.13 Affinity of overlapping 15-mers spanning the C-peptide:WE14 polypeptide. Binding of twenty-seven overlapping 15-mer peptides (labelled 1-27; the first amino acid of the 15-mer is indicated directly below the peptide number) to HLA-DR4, -DQ8, -DR3 and -DQ2 was predicted using the NetMHCII 2.3 algorithm to compute a percentile rank (% rank). Overlapping peptides were offset by one residue. The dotted lines on the graph at 2% and 10% indicate the threshold for a strong and weak binder, respectively. Bold indicates the first amino acid of the 15-mer where the core sequence contains the hybrid junction. For HLA-DR4, the binding affinity of R5 (IA-2 709-736; LAKEWQALCAYQAEPNTCATAQGEGNIK) was computed as a positive control.

The above data suggested that a hybrid epitope capable of binding with sufficient affinity to the high-risk HLA class II alleles does not exist within the C-peptide:WE14 polypeptide. However, this observation depends on the algorithms ability to accurately predict peptide

binding. To determine this, the same algorithm was used to predict the binding affinity of eluted peptides for their corresponding HLA allele. A % rank was generated for each eluted peptide by rounding to the nearest integer and percentile rank plotted as a function of peptide frequency (Figure 4.14A). Eluted peptides were predicted to have a wide range of affinities indicating that the approach does not bias towards identifying peptides with only the highest affinity for HLA. However, the peptide frequency diminished with increasing percentile rank for all HLA alleles. Although, this was less pronounced for HLA-DQ2 where the median percentile rank (17%) was greater than that of HLA-DR4 (6.5%), -DQ8 (10%) and -DR3 (3.5%). These data also suggest that the HLA-DR-peptide repertoire is skewed toward a higher affinity than the DQ repertoire as a distinct peak in peptide frequency (>10%) is observed at a percentile rank of 1%. This, however, may be a feature of the binding algorithm in that larger datasets are available on DR-eluted peptides compared to DQ, permitting identification of more stringent binding rules and therefore more accurate binding prediction.

The frequency of eluted peptides characterised as strong and weak binders was also analysed, with peptides considered to be non-binders (% rank ≥10) categorised further into four groups according to their % rank (Figure 4.14B). Overall, strong binder peptide frequencies ranged from 11% to 25% and weak peptide binder frequencies from 25% to 33%. For HLA-DR4, 58% of eluted peptides were predicted to be binders, having a % rank <10. Of these, 25% were predicted to be strong binders and 33% predicted to be weak binders. The greatest proportion of non-binders was represented by peptides with a % rank less than 25% and this was consistent across all the alleles. Similar data was obtained for HLA-DQ8 with approximately 50% of eluted peptides predicted to bind. 76% of peptides eluted from HLA-DR3 were predicted to bind with a near equal distribution of weak and strong binders. The lowest frequency of binders (36%) was seen by analysis of peptides eluted from HLA-DQ2, suggesting *in silico* prediction

of binding to -DQ2 is currently lacking. To summarise, these data indicate peptides that are predicted to bind and those that are eluted represent two different entities, as many eluted peptides are not predicted to bind their cognate class II molecule. The poor binding prediction results of the C-peptide:WE14 polypeptide do not therefore preclude elution of a hybrid epitope.

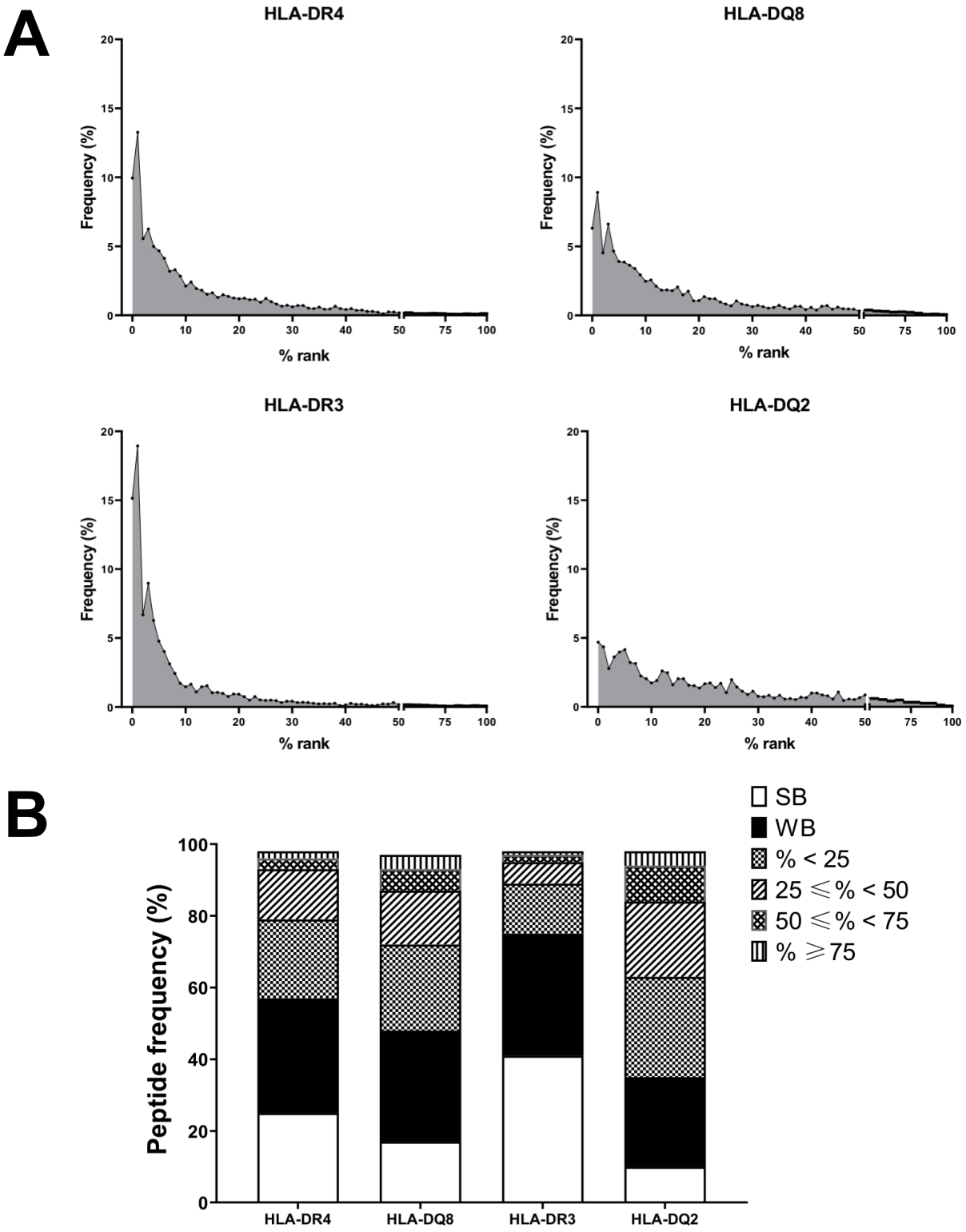


Figure 4.14 Affinity of eluted peptides. The % rank of eluted peptides was predicted using NetMHCII 2.3 and plotted against peptide frequency for each HLA molecule (A). B shows the percentage of peptides characterised as a strong binder (% rank <2), weak binder (% rank <10) or a non-binder (% rank ≥10; peptides falling into this category are divided into four groups indicated by patterned bars) to a given HLA allele. In this case, predictions were performed for the six most prevalent peptide lengths for each HLA molecule. HLA, human leukocyte antigen; SB, strong binder; WB, weak binder.

4.11.2 The hybrid polypeptide is not efficiently processed

A second plausible explanation for not eluting a hybrid epitope is that the polypeptide species cannot be processed to generate an epitope sequence containing the hybrid junction. Using the program SitePrediction (Verspurten et al., 2009) the susceptibility of the C-peptide:WE14 polypeptide to proteolytic cleavage was assessed. Cathepsins S, B, D and L were chosen as putative enzymes as they represent lysosomal proteases involved in class II antigen processing expressed by Preiss EBV-transformed B cells (Wang et al., 2009). As a control, predicted cleavage sites within a sequence of amino acids from GAD65(200-300) containing the immunodominant 274-286 epitope (IAFTSEHSHFSLK) were also determined (Figure 4.15).

The predicted cleavage sites within the GAD260-300 sequence occur around the 274-286 epitope region but never within it. In the case of C-peptide:WE14, a predicted cleavage site for cathepsin D lies at the junction between the C-peptide and WE14 components, thus, cleavage at this position would likely destroy the hybrid epitope. Cleavage sites for cathepsins B, L and S lie exclusively in the C-peptide region of the hybrid polypeptide. Cleavage by any one of these enzymes at the site proximal to the intersection of the two peptide species, gives rise to peptides of between twenty-two and twenty-four residues derived from both C-peptide and WE14. Such peptides will still be linked at their C-terminus to PWM as cleavage sites within the WE14 sequence by cathepsins S, B, D and L are not predicted to exist. Therefore, it is reasonable to envisage that the cleavage products may require C-terminal trimming in order to be available for epitope generation. Such results do not exclude elution of a hybrid epitope as the DLQTLAL*WSRM* sequence has been identified in granule-enriched fractions of mouse β-cell extracts further processed by cleavage with an endoproteinase (Delong et al., 2016). Consequently, a cleavage or trimming site within WE14 must exist. It is feasible that an enzyme specific to the β-cell granule is capable of cleaving chromogranin A to generate WE14. It is

important to note, given the potential generation of peptide species >20 amino acids in length, that the full polypeptide sequence was used for PEAKS database searching allowing for any peptide derived from this sequence to be identified irrespective of length. Longer peptides are generally more difficult to identify by MS however, peptides up to 45 amino acids were detected in the current elution dataset. Overall, the data supports the idea that the hybrid polypeptide is not efficiently processed to generate a hybrid epitope, largely due to a lack of predicted cleavage sites within the WE14 sequence.

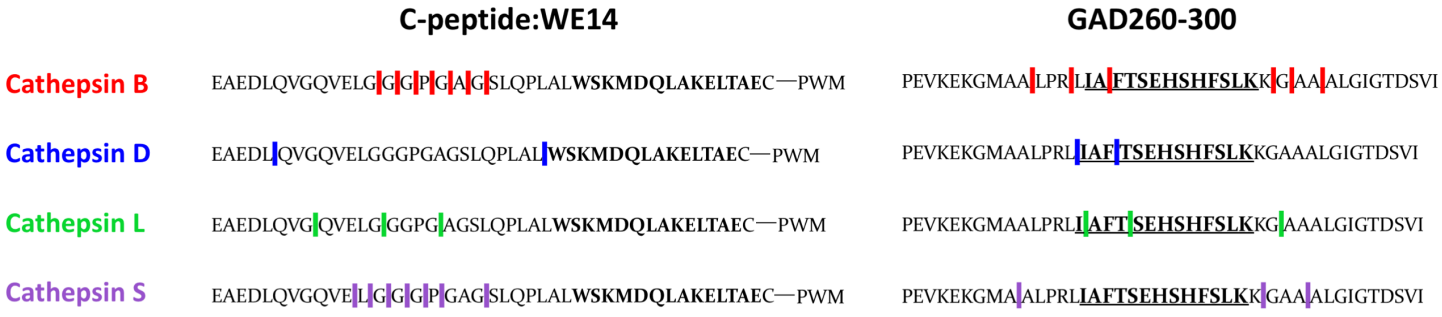


Figure 4.15 Susceptibility of C-peptide:WE14 and GAD260-300 sequences to lysosomal degradation by cathepsins. Figure shows predicted cleavage sites by cathepsins B, D, L and S indicated by coloured vertical bars within the sequence using the program SitePrediction. Only cleavage sites with a specificity greater than 95% are shown. Specificity represents the probability that a site is a bona fide cleavage site. Amino acids displayed in bold indicate the WE14 component of C-peptide:WE14 (left) and the immunodominant 274-286 epitope (right; also underlined). PWM, pokeweed mitogen; GAD260-300, amino acids 260-300 from glutamic acid decarboxylase (GAD65).

4.11.3 Immunological versus β-cell processing

It is also possible that the polypeptide species does not require immunological processing but processing by the β-cell. Recently, β-cells have been described to have the capacity to secrete peptides into the periphery via exocytosis of crinophagic granules (Wan et al., 2018) which may or may not require further processing.

4.12 Concluding remarks

In this chapter, novel hybrid epitopes could not be identified by elution of peptide-HLA complexes from EBV B cells pulsed with an ADS-delivered hybrid polypeptide. However, identification of a pokeweed-derived epitope, nested peptide sets and findings consistent with previously reported class II elution data, for example peptide length and binding motif, provided evidence of technical success. Later in the chapter, I discuss the possible reasons for not being able to elute a hybrid epitope. These include inefficient processing, binding and/or the existence of an alternative mechanism of hybrid peptide generation.

CHAPTER 5: IMMUNOGENICITY OF HYBRID POLYPEPTIDES

5.1 Background to the Chapter

One of the conclusions drawn from the previous chapter was that naturally processed and presented epitopes derived from the C-peptide:WE14 polypeptide are not identified by peptide elution of ADS-pulsed EBV B cells. Following this, one could hypothesise that hybrid polypeptides are not immunogenic. This chapter aims to explore the immunogenicity of hybrid polypeptides at the single-cell level by FluoroSpot. Previous literature describing HIP-reactive T cells to have a pro-inflammatory phenotype (Baker et al., 2019a; Baker et al., 2018; Baker et al., 2019b) prompted me to use a triple-colour assay format, permitting detection of three pro-inflammatory cytokines: IFN-γ and the Th17 cytokines IL-17A and IL-22 (Liang et al., 2006).

Enzyme-linked immunospot (ELISpot) and FluoroSpot assays are instrumental technologies that have allowed detection of $CD4^+$ T cell responses to native islet autoantigen peptides in peripheral blood; with the quality of immune response shown to be different in patients with T1D and in HLA-matched controls (Arif et al., 2004). In the case of hybrid peptides, Baker and colleagues (2019) demonstrated significantly elevated responses to several HIPs in the peripheral blood of T1D patients by IFN-γ ELISpot. In contrast, nondiabetic controls did not show significantly increased responses to any HIPs, highlighting that T cell responses to hybrid peptides are disease-relevant. An assumption of both studies is that the length of the peptides tested are such that they bind externally to HLA molecules without processing. Since ADS labelling is not 100% efficient allowing for some free peptide to be present during pulsing, it is feasible to presume the C-peptide:WE14 polypeptide does not bind HLA directly or else this species (or fragments of the polypeptide) would have likely been identified by peptide elution. Thus, the main aim of the present study was to evaluate the ability of hybrid polypeptides to

be processed and presented by APCs present in cryopreserved peripheral blood mononuclear cell (PBMC) samples from newly diagnosed T1D patients (n=20). Control subjects were not tested as this piece of work aimed to address immunogenicity, not whether a difference between patients with T1D and healthy controls exists. Subjects included in this study are described in detail in Table 2.4 (Materials and Methods).

5.2 Design of peptide panel

Processing of the long C-peptide:WE14 polypeptide was investigated by comparing the response to that of a short (14-mer) pre-processed epitope (hEL:ChgA-WE14) shown in Table 2.2 (Materials and Methods). Of note, the sequence of the hEL:ChgA-WE14 epitope is contained within the long C-peptide:WE14 polypeptide. A secondary aim was to consider the effect of reversing the orientation of the peptide species within the polypeptide, i.e. instead of C-peptide being present at the N-terminus of the polypeptide, it exists at the C-terminus (WE14:C-peptide). In this case, the short hEL:ChgA-WE14 epitope is absent from the reversed orientation polypeptide. A pre-processed epitope and reversed orientation polypeptide were also designed for the C-peptide:IAPP2 polypeptide (termed hEL:IAPP2 and IAPP2:C-peptide) and are indicated in Table 2.2.

5.3 Overview of FluoroSpot analysis

Figure 5.1A presents representative FluoroSpot well images of T cell responses to a long and short peptide shown for one patient. Using a cut-off (as defined in Materials and Methods), positive responses were identified with one patient (NDB121) found to elicit a response to all peptides within the panel with an unusually high magnitude (stimulation index, SI >100) for IFN-γ specifically (Figure 5.1B). This is suspected to be a spurious result as autoreactive responses do not occur to this level. Subsequently, this patient was excluded from downstream

analysis. An overview of the cytokine responses is provided (Figure 5.1C) with specific responses being teased out in the succeeding figures. The triple-colour assay determined the simultaneous production of IFN-γ, IL-17A and IL-22, allowing for double- and triple-secreting cells to be enumerated. Dual-secreting cells were sparse following peptide stimulation with often only one individual eliciting a positive response to a given peptide. In these rare cases, the response was most commonly directed against a short hybrid epitope. Dual and triple-secretors were present at a far greater frequency in cells stimulated with Infanrix-hexa and *Candida* antigen (positive control).

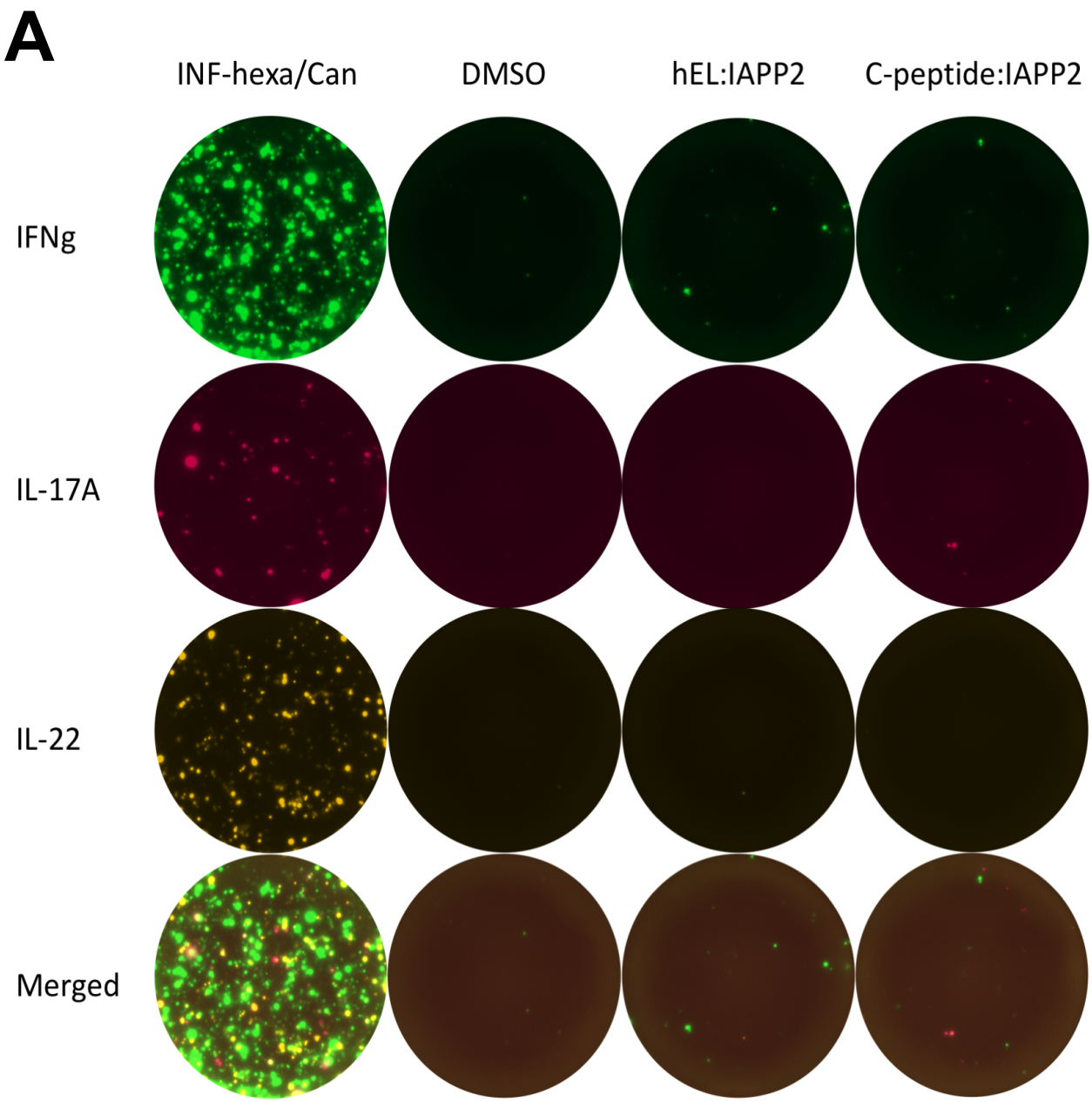


Figure 5.1 T cell responses to hybrid (poly)peptides. A triple-colour FluoroSpot analysing the production of IFN-γ (green spots), IL-17A (red spots) and IL-22 (yellow spots) from T1D-derived PBMCs in response to peptide/polypeptide stimulation. Representative well images are shown for one patient (NDB118) following stimulation with Infanrix-hexa and *Candida* antigen, DMSO peptide diluent, hEL:IAPP2 and C-peptide:IAPP2 (A). The stimulation index (SI) for each peptide response per subject is plotted in B. Filled circles represent a positive response (SI ≥3; threshold indicated by dotted line) and open circles denote a negative response. Data shown for INF-hexa/Can (black) and hybrid (poly)peptides containing WE14 (blue) or IAPP2 (red). Tringles indicate patient NDB121 considered to be an outlier. Excluding NDB121 from analysis, the frequency of IFN-γ (red), IL-17A (blue) and IL-22 (green) responders for each (poly)peptide was determined (C). INF-hexa/Can, Infanrix-hexa and *Candida* antigen.

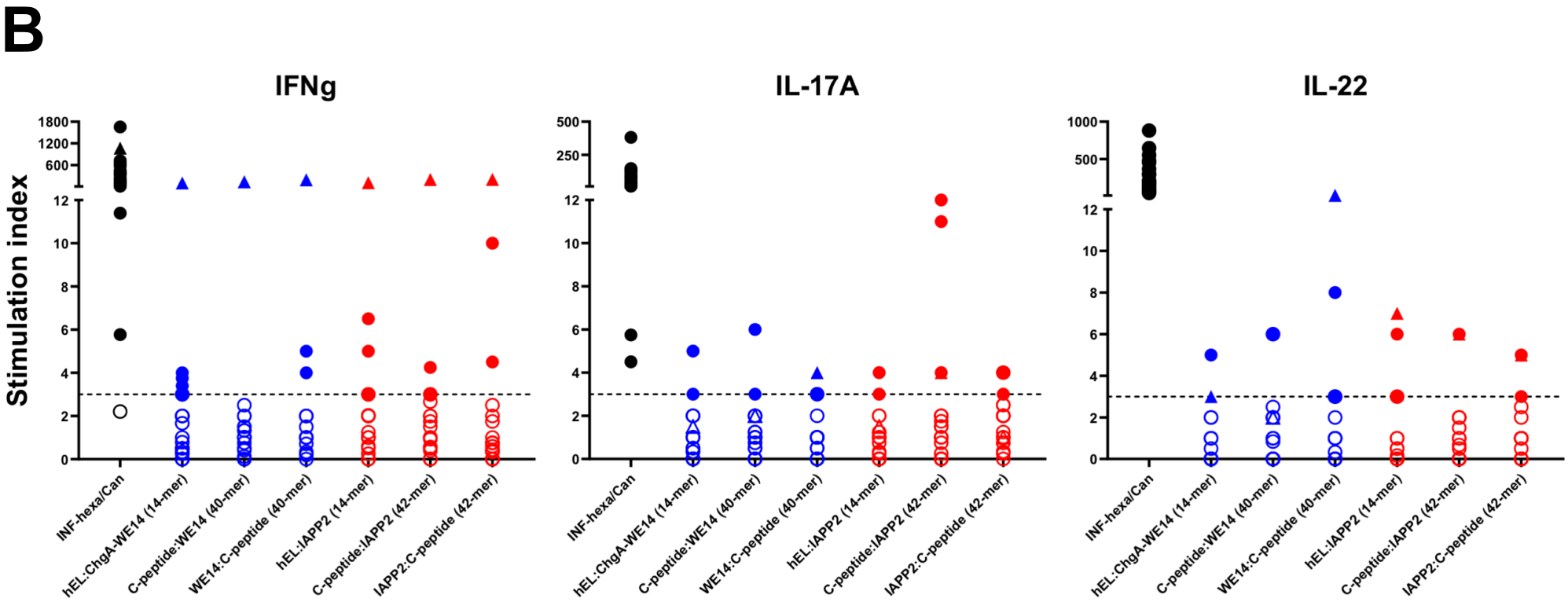


Figure 5.1 T cell responses to hybrid (poly)peptides. A triple-colour FluoroSpot analysing the production of IFN-γ (green spots), IL-17A (red spots) and IL-22 (yellow spots) from T1D-derived PBMCs in response to peptide/polypeptide stimulation. Representative well images are shown for one patient (NDB118) following stimulation with Infanrix-hexa and *Candida* antigen, DMSO peptide diluent, hEL:IAPP2 and C-peptide:IAPP2 (A). The stimulation index (SI) for each peptide response per subject is plotted in B. Filled circles represent a positive response (SI ≥3; threshold indicated by dotted line) and open circles denote a negative response. Data shown for INF-hexa/Can (black) and hybrid (poly)peptides containing WE14 (blue) or IAPP2 (red). Tringles indicate patient NDB121 considered to be an outlier. Excluding NDB121 from analysis, the frequency of IFN-γ (red), IL-17A (blue) and IL-22 (green) responders for each (poly)peptide was determined (C). INF-hexa/Can, Infanrix-hexa and *Candida* antigen.

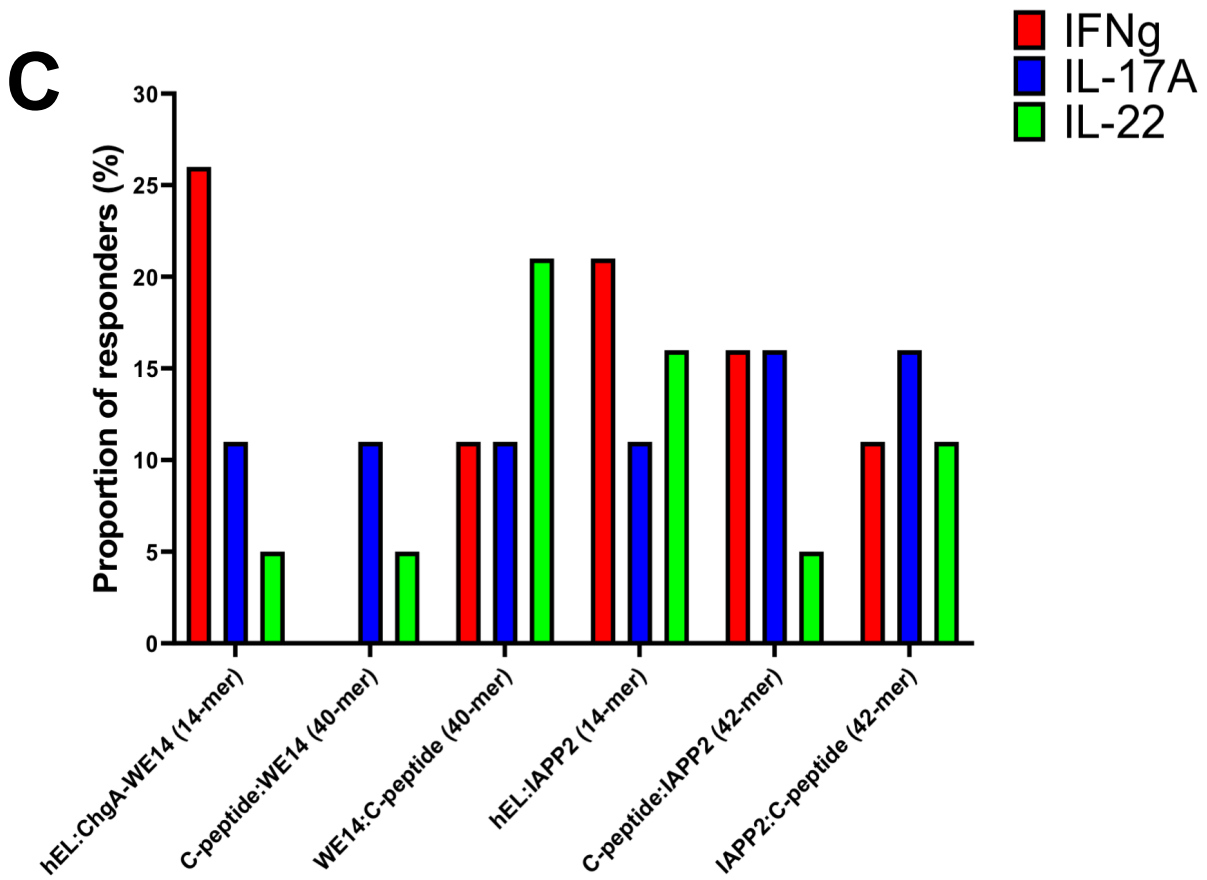


Figure 5.1 T cell responses to hybrid (poly)peptides. A triple-colour FluoroSpot analysing the production of IFN-γ (green spots), IL-17A (red spots) and IL-22 (yellow spots) from T1D-derived PBMCs in response to peptide/polypeptide stimulation. Representative well images are shown for one patient (NDB118) following stimulation with Infanrix-hexa and *Candida* antigen, DMSO peptide diluent, hEL:IAPP2 and C-peptide:IAPP2 (A). The stimulation index (SI) for each peptide response per subject is plotted in B. Filled circles represent a positive response (SI ≥3; threshold indicated by dotted line) and open circles denote a negative response. Data shown for INF-hexa/Can (black) and hybrid (poly)peptides containing WE14 (blue) or IAPP2 (red). Tringles indicate patient NDB121 considered to be an outlier. Excluding NDB121 from analysis, the frequency of IFN-γ (red), IL-17A (blue) and IL-22 (green) responders for each (poly)peptide was determined (C). INF-hexa/Can, Infanrix-hexa and *Candida* antigen.

5.4 Processing of hybrid polypeptides

To address the question whether hybrid polypeptides receive immunological processing, the magnitude and frequency of the polypeptide response was compared to that of the corresponding short pre-processed epitope. IFN-γ responses to the short 14-mer epitopes exhibited a trend toward higher stimulation indices than the polypeptides, particularly discernible between C-peptide:WE14 and hEL:ChgA-WE14, although this was not statistically significant (Figure 5.2A). A greater prevalence of response to the short epitopes was also observed and this reached statistical significance for hEL:ChgA-WE14 ($p=0.0463$) where no patients elicited a positive IFN-γ response to C-peptide:WE14 (Figure 5.2B). 16% (3/19) of individuals triggered a response characterised by IFN-γ secretion against C-peptide:IAPP2 and this increased to 21% (4/19) following stimulation with the hEL:IAPP2 epitope. The same trends were not however, replicated in the case of IL-17A and IL-22 (Figure 1C). Altogether, these data indicate that the short epitopes are immunogenic since they do not require processing but bind HLA externally. The polypeptide which contains the short epitope, requires processing and this likely results in the generation of fewer peptides, thus providing an explanation for the reduced prevalence and magnitude of response. The lack of response to C-peptide:WE14 does however suggest either limited processing of this particular species and/or T cells are non-responsive to the epitopes that are generated. In contrast, C-peptide:IAPP2 is seemingly more available for epitope generation although the response is not necessarily the result of T cells recognising both long and short peptides.

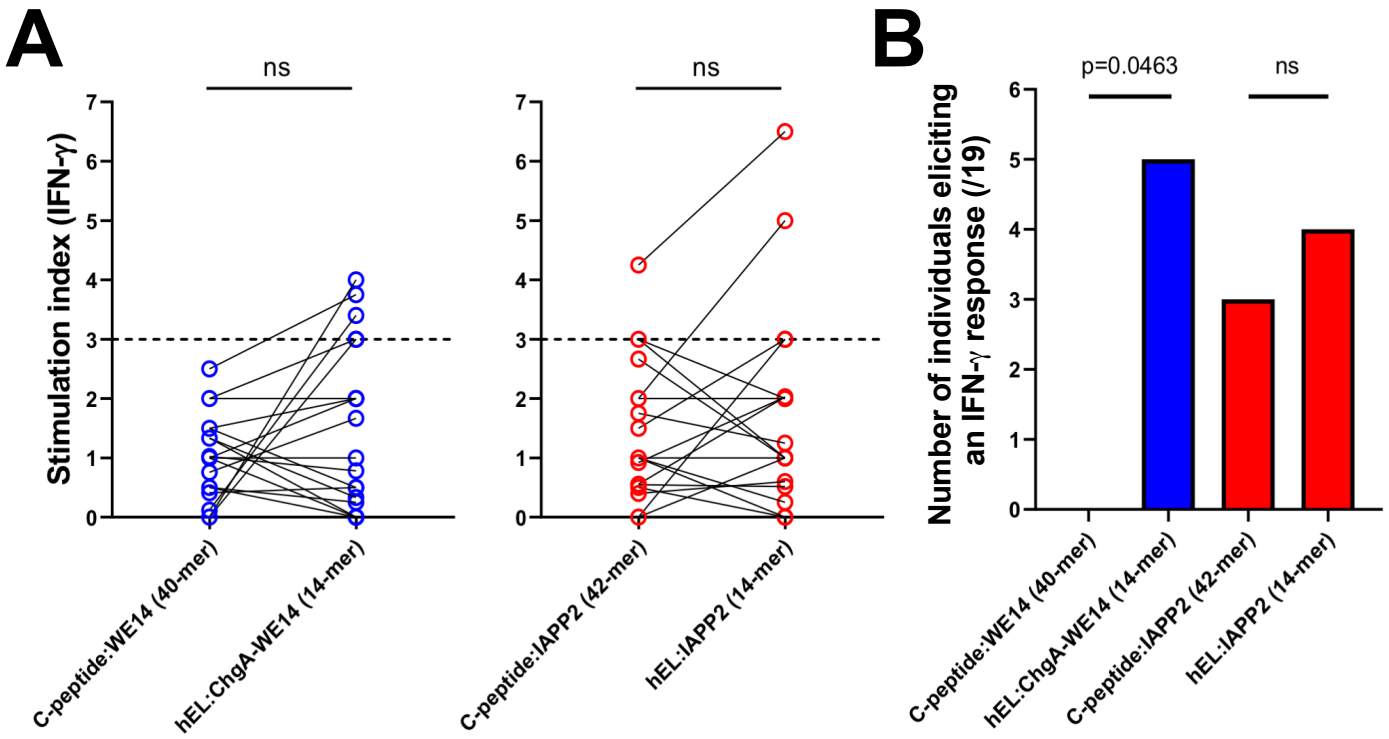


Figure 5.2 Pre-processed hybrid epitopes elicit a trend toward a greater prevalence and magnitude of response. The stimulation index (SI) of responses to the hybrid polypeptides and their corresponding 14-mer epitopes has been plotted for IFN-γ (A). The dotted line on the graph indicates the threshold for a responder (SI ≥3). Paired data was tested for significance using Wilcoxon matched pairs test. The number of individuals eliciting a positive IFN-γ response is indicated in B and a p-value calculated using Fisher's exact test. ns; not significant.

5.5 Reversing the orientation of peptides within the hybrid polypeptide

To next examine the effect of reversing the orientation of the hybrid polypeptide, the frequency of responders to the N-terminal C-peptide polypeptide was compared to that of the reversed orientation polypeptide (Figure 5.3). The data show that for IFN-γ and IL-22, the number of individuals eliciting a positive response was greater following stimulation of PBMCs with WE14:C-peptide than with C-peptide:WE14 (Figure 5.3A). For IL-17A, stimulation with either the normal or reversed orientation polypeptide triggered an equal number of responders for both WE14- (2/19) and IAPP2- (3/19) containing polypeptides. Given that the three cytokines tested all represent pro-inflammatory cytokines, total responders can also be identified based on the number of individuals eliciting a positive response irrespective of the

cytokine (Figure 5.3B). In this case, 16% (3/19) of individuals elicited a cytokine response to C-peptide:WE14 and this frequency doubled following stimulation with WE14:C-peptide. An equal number of total responders (5/19) was observed for C-peptide:IAPP2 and IAPP2:C-peptide. Taking all the cytokine responses together, the data provide evidence that long hybrid polypeptides can be immunogenic particularly in the reverse orientation.

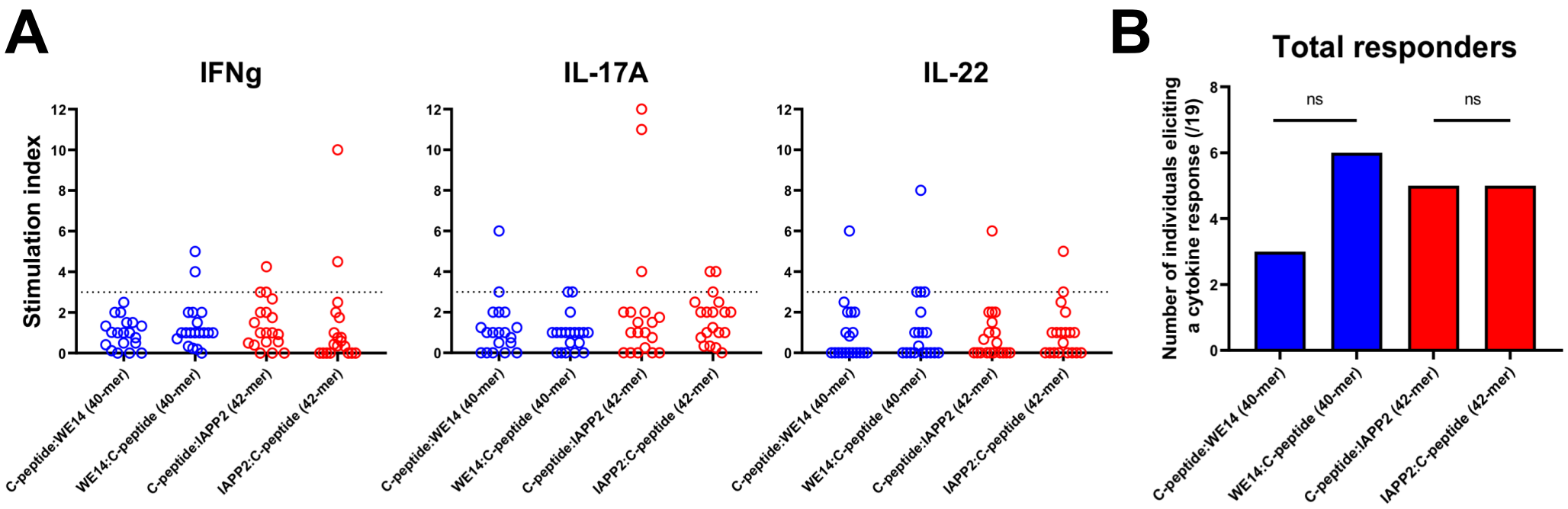


Figure 5.3 Greater frequency of responders to WE14:C-peptide than to C-peptide:WE14. The stimulation index (SI) of responses to the normal and reversed orientation hybrid polypeptide has been plotted for all of the cytokines tested (A). The dotted line on the graph indicates the threshold for a responder (SI ≥3). The number of individuals eliciting a positive cytokine response is indicated in B and a p-value calculated using Fisher's exact test. ns; not significant.

5.6 Assessing T cell responses to left and right peptides

The presence or absence of shared responses was used to evaluate whether T cell responses to the normal and reversed hybrid polypeptides were directed against common regions within the sequence e.g. C-peptide (Figure 5.4). In most individuals the response to either the normal or reversed orientation polypeptide was dominated by one cytokine, with the production of IFN-γ not necessarily accompanied by the production of IL-17A and/or IL-22. No shared responses were observed if the hybrid polypeptide contained WE14, however two individuals (G832 and NDB118) elicited a shared response to IAPP2-containing hybrid polypeptides. G832 mounts a positive IL-17A response to both the normal (C-peptide:IAPP2) and reversed (IAPP2:C-peptide) polypeptide, whereas NDB118 responds to the N-terminal C-peptide polypeptide to produce IL-17A and IFN-γ. In the reversed orientation, NDB118 elicited a positive response for all three cytokines. The rarity of shared responses between normal and reversed hybrid polypeptides, formally shown by a lack of statistical significance (C-peptide:WE14 versus WE14:C-peptide, $p=0.5170$; C-peptide:IAPP2 versus IAPP2-C-peptide, $p=0.5696$) suggests that the T cell response is not directed against the C-peptide or WE14/IAPP2 region but likely against peptides being exclusive to either of these hybrid polypeptides. It could therefore be speculated that in both cases responses are directed against an epitope that contains residues of the hybrid junction.

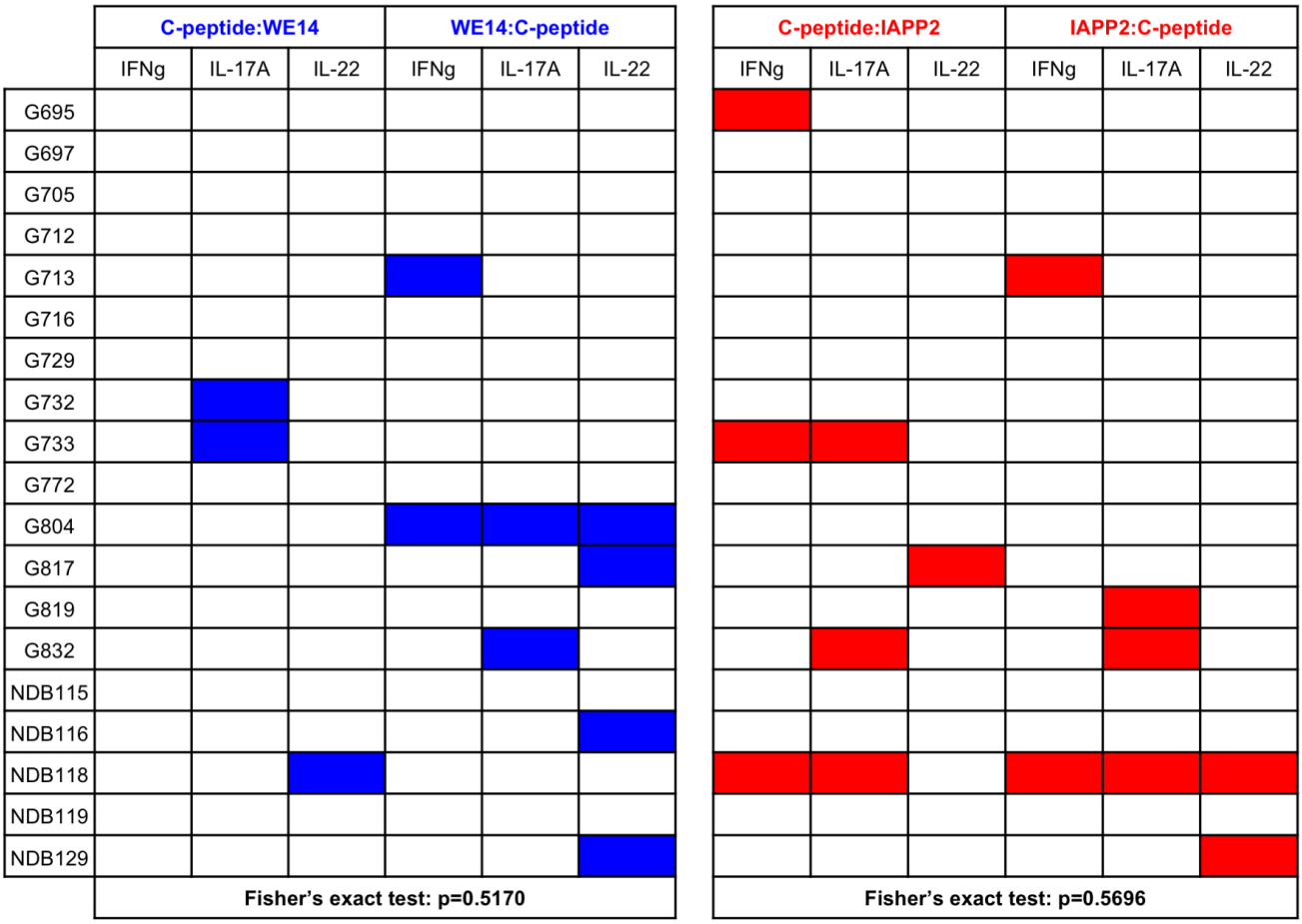


Figure 5.4 Shared responses between N-terminal C-peptide and reversed orientation hybrid polypeptides are rare. PBMCs from nineteen T1D patients were stimulated with either a N-terminal C-peptide or reversed orientation hybrid polypeptide. WE14- and IAPP2-containing polypeptides are indicated by blue and red, respectively. Filled rectangles represent a positive response (SI ≥3) measured through production of IFN-γ, IL-17A and IL-22. A p-value was calculated (Fisher's exact test) for comparing responses to the two long peptides using a 2x2 contingency table containing a matrix of responders and non-responders. A responder was defined by a SI ≥3 for at least one of the cytokines. SI, stimulation index.

5.7 Evaluating processing of a polypeptide as a mechanism of hybrid epitope generation

Shared responses between the long hybrid polypeptide and the short hybrid epitope can be an indication of whether processing of the hybrid polypeptide generates the hybrid epitope (Figure 5.5). If the hybrid epitope is indeed generated, where a response to the shortened peptide epitope is observed one would also expect to see a response to the long polypeptide. 60% (3/5) of patients responding to hEL:ChgA-WE14 elicit a cytokine response which is exclusive to the

short epitope i.e. they do not mount a response to C-peptide:WE14. Similarly, 63% (5/8) of individuals triggered a positive cytokine response to hEL:IAPP2 but this did not coincide with a response to C-peptide:IAPP2. The lack of statistical significance between the two groups (C-peptide:WE14 versus hEL:ChgA-WE14, p=0.1548; C-peptide:IAPP2 versus hEL:IAPP2, p=0.6027) provides further confidence in deducing that a response to the hybrid polypeptide does not always parallel a response to the short hybrid epitope. A reasonable interpretation of this is that immunological processing and presentation of a hybrid polypeptide to generate the short hybrid epitope is insufficient which does not reach the threshold for T cell activation. Notably, a few patients (G732, G695 and G733) respond to the long polypeptide but not the short epitope. In this case, processing of the hybrid polypeptide may generate non-hybrid epitopes which are targeted or other epitopes wherein the hybrid contributes but in a different register than the short epitope.

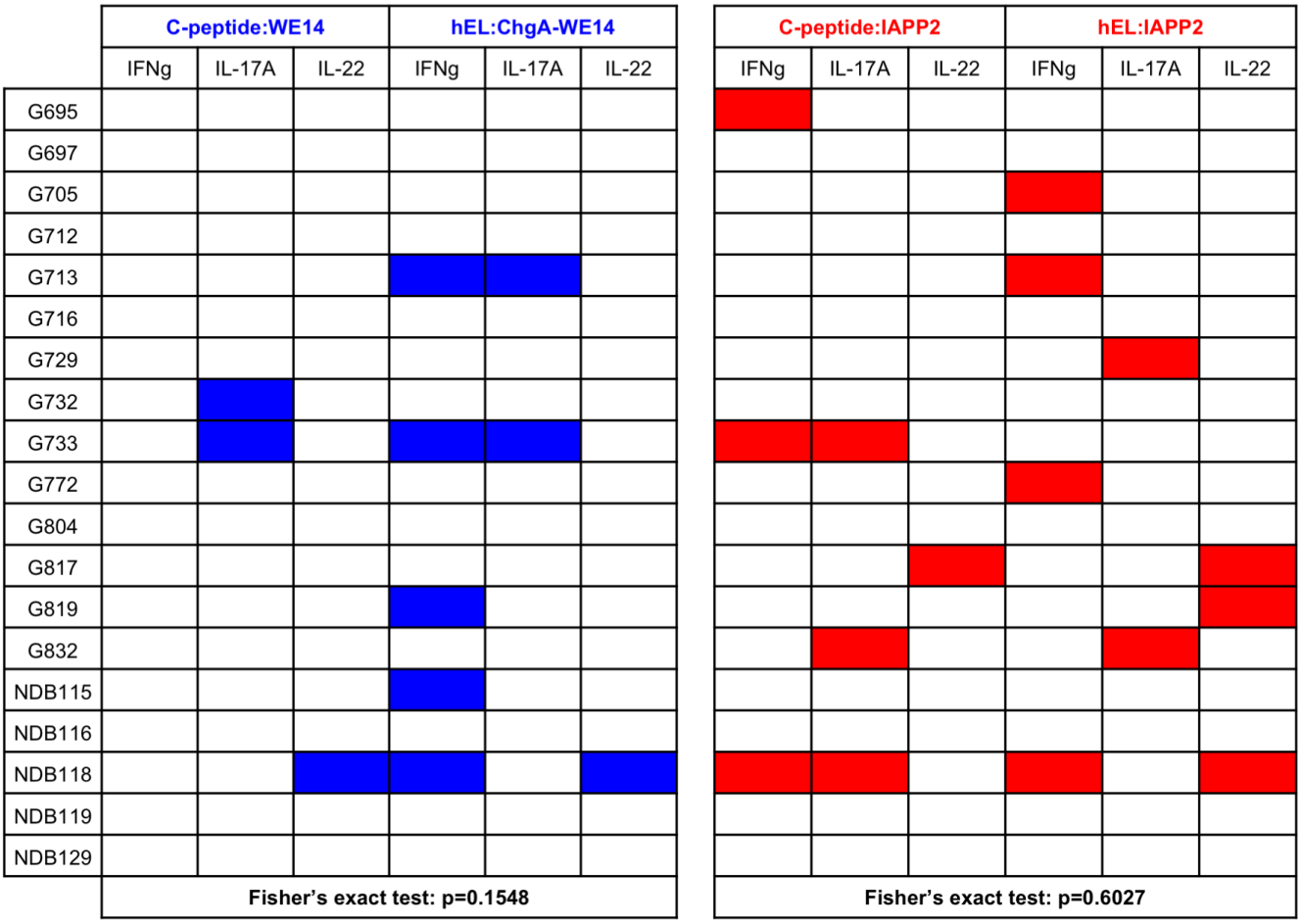


Figure 5.5 A response to the hybrid polypeptide does not always coincide with a response to the short hybrid epitope. PBMCs from nineteen T1D patients were stimulated with either a hybrid polypeptide or a short hybrid epitope. WE14- and IAPP2-containing peptides are indicated by blue and red, respectively. Filled rectangles represent a positive response (SI ≥3) measured through production of IFN-γ, IL-17A and IL-22. A p-value was calculated (Fisher's exact test) for comparing responses to the short and long peptide using a 2x2 contingency table containing a matrix of responders and non-responders. A responder was defined by a SI ≥3 for at least one of the cytokines. SI, stimulation index.

5.8 HLA-unrestricted responses to hybrid species

A tempting notion is that patients responding to the long polypeptide but not the short, bear a distinctive HLA haplotype. This was investigated by correlating the cytokine response to the hybrid polypeptide (normal and reversed) and the shorter hybrid epitope with the HLA haplotype of the patient cohort (Figure 5.6). The HLA genotype of patients eliciting a response

to the long but not the short hybrid epitope was also found to be present in individuals responding to both peptides. Similarly, no clear correlation with HLA haplotype was observed for individuals responding exclusively to either the normal or reversed hybrid polypeptide. The absence of any correlation is likely because the dataset is of insufficient size, making responses few and far between. Any possible HLA restriction could be more accurately determined using tetramers loaded with hybrid polypeptides. Bearing in mind that Delong et al. (2016) reported responses to hybrid peptides were HLA-DQ8-restricted, it was however interesting to find in the current study that responses also occurred in non-DQ8 individuals. For example, five patients responded to hEL:ChgA-WE14, three of which had -DQ2 haplotypes, which might be explained in part by the shared features of HLA-DQ2 and -DQ8 binding motifs (Godkin et al., 1997). Overall, the data suggest the existence of additional factors bedsides HLA haplotype which determine whether a T cell response is elicited to a hybrid polypeptide and/or a hybrid epitope.

Figure 5.6 Responses are not limited to individuals carrying the HLA-DQ8 haplotype. PBMCs from nineteen T1D patients were stimulated with either a N-terminal C-peptide polypeptide, reversed orientation polypeptide or a short hybrid epitope. WE14- and IAPP2-containing peptides are indicated by blue and red, respectively. Filled rectangles represent a positive cytokine response defined by a SI ≥3 for at least one of the cytokines: IFN-γ, IL-17A and IL-22. The HLA-DR and -DQ haplotype of each patient is indicated by the black squares. SI, stimulation index.

5.9 Binding and core sequence prediction of reversed hybrid polypeptides

A yet unexplored finding of the current study is the increased frequency of responders to the WE14:C-peptide polypeptide compared with the N-terminal C-peptide polypeptide (C-peptide:WE14). To explore why this might be the case, overlapping 15-mers (offset by one residue) of the normal and reversed orientation polypeptides were designed and their binding affinity to HLA-DQ8 and -DQ2 assessed using the NetMHCII 2.3 algorithm. The algorithm identifies strong and weak binders with a percentile rank of <2 and <10, respectively. Non-

binders are assigned a percentile rank of ≥10. Of the twenty-six overlapping peptides of C-peptide:WE14, none were predicted to be strong binders to HLA-DQ2 whereas four overlapping peptides of WE14:C-peptide were predicted to bind with high affinity (Figure 5.7). Similarly, peptides spanning C-peptide:IAPP2 contained no strong binders to HLA-DQ2 while five of the IAPP2:C-peptide overlapping peptides were predicted to be strong binders. This observation was also true for binding to HLA-DQ8, where peptides overlapping the reversed orientation polypeptide contained a larger number of strong (and weak) binders than the N-terminal C-peptide polypeptide. Altogether, the reversed hybrid polypeptides were predicted to harbour a greater number of candidate binders to T1D-associated HLA-DQ alleles than N-terminal C-peptide hybrid polypeptides. These findings may, in part, explain the increased frequency of responders to the WE14:C-peptide polypeptide.

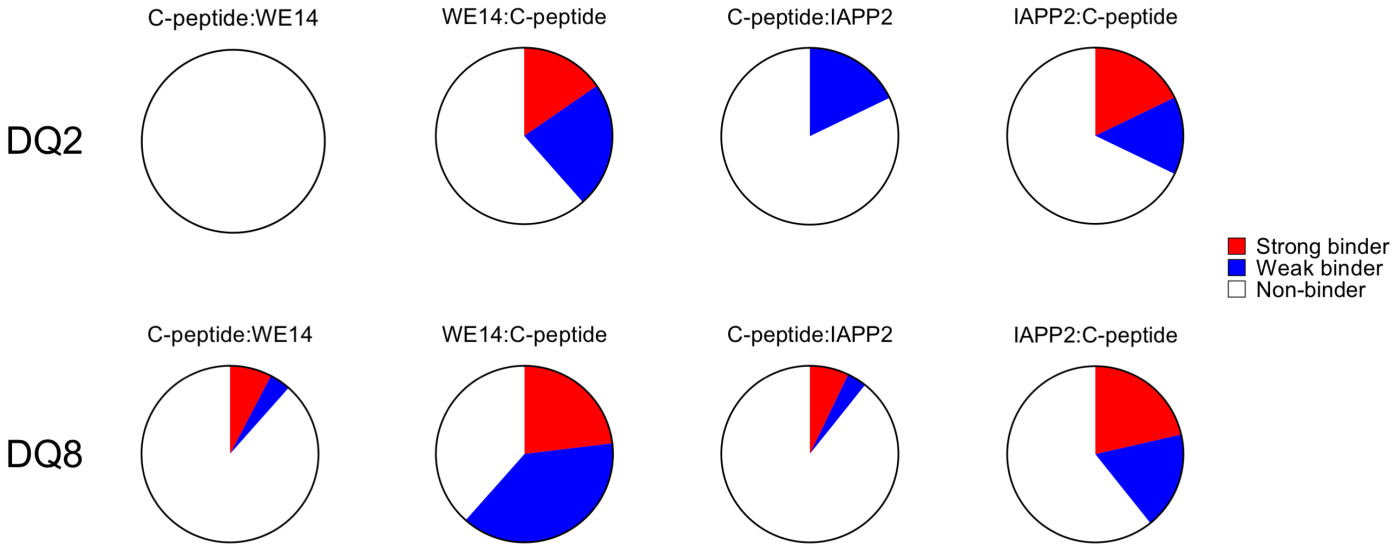


Figure 5.7 Overlapping peptides spanning the reversed hybrid polypeptide sequence contain more candidate binders to HLA-DQ2/8. Binding of overlapping 15-mer peptides spanning N-terminal C-peptide and reversed hybrid polypeptides to HLA-DQ8 (bottom) and -DQ2 (top). Binding affinity was predicted using the NetMHCII 2.3 algorithm to compute a percentile rank (% rank) and peptides characterised as strong (% rank <2, red), weak (% rank <10, blue) or non-binders (% rank ≥10, white). The pie chart reflects the relative number of peptides in each group. Overlapping peptides were offset by one residue.

To investigate whether the increased binding capacity of the reversed hybrid polypeptides is mediated by residues from both peptide species (i.e. C-peptide and WE14/IAPP2), the core binding sequence of each overlapping 15-mer was determined (Figure 5.8). The results show for HLA-DQ2, the core sequences of binders (strong and weak) exclusively contain the hybrid junction (Figure 5.8A, C). The binding core of strong binders to HLA-DQ8 incorporate only the C-peptide region however, the majority of weak -DQ8 binders were predicted to have a binding core containing residues of the hybrid junction (Figure 5.8B, D). These data indicate the enhanced binding capacity of the reversed polypeptides relative to the normal orientation, particularly for binding to HLA-DQ2, is due to the altered sequence of amino acids at the intersection of the two peptides.

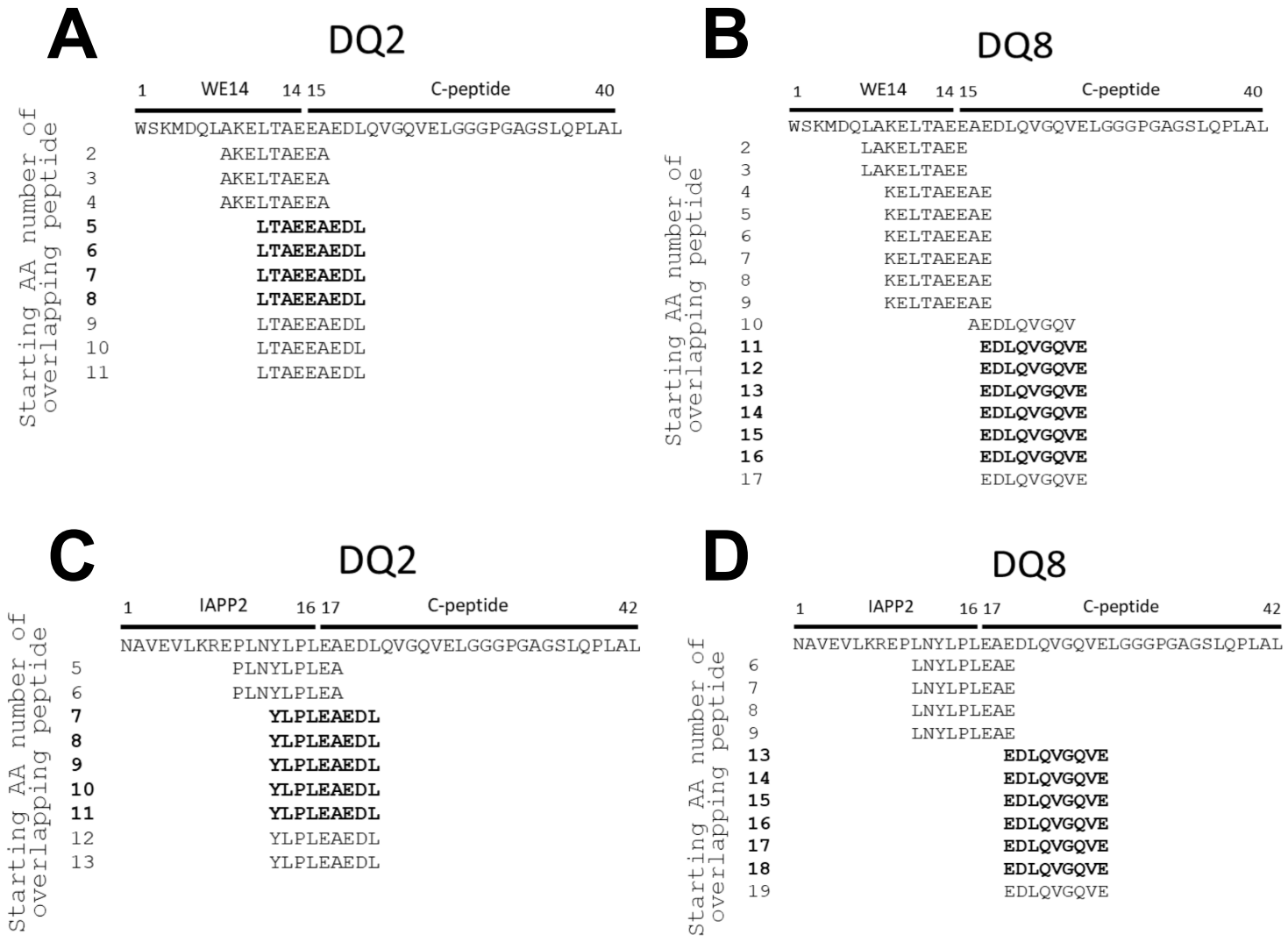


Figure 5.8 Predicted binding cores of reversed polypeptides contain the intersection of the two granule peptides. Binding cores of overlapping 15-mer peptides spanning reversed polypeptide sequences (WE14:C-peptide, top; IAPP2:C-peptide, bottom) predicted to bind HLA-DQ8 (B and D) and -DQ2 (A and C). Binding affinity (% rank) and core sequences of 15-mer peptides were predicted using the NetMHCII 2.3 algorithm. Only overlapping peptides with core sequences predicted to bind with strong or weak affinity (% rank <10) are shown. Bold indicates the core sequence of overlapping peptides predicted to bind with high affinity (% rank <2). AA, amino acid.

5.10 Concluding remarks

The data presented in this chapter suggests that hybrid polypeptides can be immunogenic, although immunological processing to generate a hybrid epitope appears limited. This may

explain why naturally processed hybrid peptides were not identified in the elution study (Chapter 4). The numerous APC populations present in PBMCs including monocytes/macrophages and DCs, may account for the discrepancy between the lack of epitopes identified by elution of polypeptide-pulsed B cells and the low-level reactivity observed in the FluoroSpot assay. Interestingly, reversing the orientation of peptides within the hybrid polypeptide renders the reversed sequence particularly immunogenic. Here I present evidence to suggest that this could be due to an enhancement of binding to HLA-DQ. It is noteworthy that not all cells following activation by their cognate peptide may release (the selected) cytokines and some may not become "functional cytokine secretors" but rather just proliferate. T cell assays based on proliferation (CFSE) or flow cytometric analysis of activation markers (e.g. CD154) may therefore provide additional information. Multimers could also be useful for characterisation of peptide-specific $CD4^+$ T cells directly *ex vivo* (Holland et al., 2015).

CHAPTER 6: DISCUSSION

6.1 Research question, methods employed and how this integrates into current knowledge

In this thesis I have investigated whether natural processing of a long hybrid polypeptide within an antigen-presenting cell is required for generation of the C-peptide:WE14 hybrid epitope. To address this question, I explored: antigen delivery, whether the epitope is naturally processed and presented and whether hybrid polypeptides can elicit immune responses. This particular polypeptide species was selected as data available in the NOD mouse prior to commencing my PhD identified a C-peptide:WE14 epitope in fractions of β-cell tumour lysates (Delong et al., 2016). T cell responses against the equivalent human epitope sequence are also present in T1D patients (Baker et al., 2019b). Moreover, unpublished data from a collaborator (Peakman & Purcell) identified a putative peptide in the human islet proteome containing C-peptide fused to WE14. I generate two lines of evidence, in an *in vitro* system and an *ex vivo* stimulation setting, to suggest that for this particular epitope a long hybrid polypeptide does not appear to be processed and presented in order to elicit a T cell response.

The research question outlined above is important because the role of hybrid peptides in T1D has attracted considerable attention over recent years. However, much of the data currently available involves unsystematic generation of hybrid peptides and identification of a corresponding T cell response, for example the insulin-insulin hybrid peptide described by Wang et al. (2019). Like many other studies, information relating to such hybrids is functional and is not based on "hard" evidence such as elution from HLA or APCs. Evidence to indicate that antigen-presenting cells are necessary for the generation and presentation of hybrid epitopes is lacking as is the nature of the precise region of the epitope(s) involved in TCR recognition. To date, evidence suggesting that hybrid peptide specific T cell clones target the intersection of the two peptides is not conclusive. Thus, the question remains whether hybrid

peptides are key targets in the pathogenesis of T1D, either eliciting a different set of T cell responses or representing an amplification of the T cell response against conventional epitopes (Harbige et al., 2017). The latter is likely to be characterised by higher peptide-HLA binding affinity of hybrid peptides over native sequences. If this is the case, it is consistent with reports that T cell responses to hybrid peptides may be stronger (e.g. cytokine production) when compared with responses to the corresponding conventional epitope (Delong et al., 2016; Wang et al., 2019).

6.2 Chapter 3: Lectin-based antigen delivery system

In the first section of this thesis I established a working antigen delivery system to identify an optimal route of antigen delivery for a polypeptide into an antigen-presenting cell. Using this antigen delivery system, which is arguably not the natural method of antigen acquisition, I explored whether a polypeptide antigen could efficiently deliver peptide epitopes for presentation in a manner that enables T cell recognition and is amenable to upscaling.

I generated two antigen delivery systems by conjugation of synthetic peptides to pokeweed mitogen or anti-DEC205 antibody. Data support that the pokeweed-based ADS method I have optimised is working as the method yielded successful delivery of the key epitope of a GAD65(274-286)-containing 41-mer peptide into EBV B cells for processing. The result is presentation of the GAD65(274-286) epitope shown using the T33 T cell hybridoma which recognises this epitope in the context of HLA-DR4 (Wicker et al., 1996). Similar experiments performed using pokeweed mitogen covalently linked to whole GAD65 antigen observed a T cell response comparable to my data and mass spectrometry also identified GAD65(274) peptides (Peakman, unpublished). Thus, the level of T cell response I observe would allow mass spectrometry to identify peptides. Delivery of the GAD65(274-286)-containing peptide

by targeting DEC205 did not yield a T cell response inconsistent with a previous report by Leung and colleagues (2013). In contrast to my study, the authors utilised whole EBV antigen which was not conjugated but rather fused to the anti-DEC05 antibody and this might explain the discrepancy. Had there been an experimental requirement, for example if pokeweed-based antigen delivery failed, it would have been useful to extend the studies with DEC205-mediated delivery further, for example by developing a system to generate a fusion antibody with the synthetic peptide attached at the C-terminus of the anti-DEC205 antibody, although such an approach would have been a challenge for upscaling.

The antigen delivery routes of pokeweed mitogen are not clearly defined as yet but may be an important determinant of peptide presentation (Cohn et al., 2013). Since pokeweed mitogen targets cell surface receptors, most likely the epitopes which are presented derive from one specific antigen processing route and this route may not result in presentation of certain epitopes. There are multiple routes by which an APC can acquire antigen, namely; macropinocytosis, receptor-mediated endocytosis, phagocytosis and autophagy (Roche and Furuta, 2015). Evidence suggests that in DCs and macrophages, the mechanism of uptake of soluble hen egg-white lysozyme (HEL) favours rapid entry into late endosomes but antigen does not remain in early endosomes for sufficient time to allow processing and cross-presentation (Belizaire and Unanue, 2009). This feature of trafficking also hindered MHC class II presentation of type B epitopes, which are generated exclusively in early endosomes (Lovitch and Unanue, 2005). These data suggest that the route of uptake can influence the epitope repertoire that is generated. Regarding the route employed by a pokeweed-based ADS, and how relevant the downstream peptide repertoire is compared to DCs (considered to be the most efficient professional APC (Guéry and Adorini, 1995)]) operating in the inflammatory environment of the pancreas or related lymph nodes is unknown.

As I have already alluded to, the mode of antigen delivery dictates trafficking through different intracellular compartments. Pokeweed mitogen binds preferentially to carbohydrate moieties on cell surface receptors with immunoglobulin-like domains (Chilson and Kelly-Chilson, 1989), consequently binding a multitude of different membrane proteins on the APC. These membrane receptors enter the endocytic recycling pathway culminating in the movement of the cargo from the early endosome to the endocytic recycling compartment (ERC) (Grant and Donaldson, 2009). From the ERC the cargo is delivered back to the cell surface via recycling endosomes. Only some of the cargo is selected in the early endosome to be delivered onto late endosomes. Whereas, in the case of phagocytic uptake, a nascent phagosome sequentially fuses with early (sorting) endosomes, late endosomes, and eventually lysosomes to form a phagolysosome (Gordon, 2016). Despite the shared involvement of the late endosome in the receptor recycling and phagocytic pathways, in general cargos in the phagocytic pathway are trafficked to different organelles than normal receptor recycling. It is therefore possible that the ADS could be diverting some of the polypeptide cargo to other pathways and not entering the antigen presentation route.

Restricting antigen delivery to a specific APC surface molecule can favour delivery into compartments equipped for processing and peptide loading. If one would have targeted only the BCR (Stevens and Peakman, 1998) this would have given confidence that the system is antigen specific however, it is known that following receptor ligation a proportion of BCRs are selectively retained at the cell surface (Hou et al., 2006) and do not feed into the same route which is favouring antigen presentation. Alternatively, complement receptor-targeted compartments also contain MHC II for peptide loading, albeit less than BCR-targeted compartments (Perrin-Cocon et al., 2004). Furthermore, efficient antigen presentation requires B cell activation (Kakiuchi et al., 1983) and although EBV infection induces expression of B

cell activation markers (Calender et al., 1987), it is not clear whether the activation status of EBV-immortalised B cells is optimal for presentation.

6.3 Chapter 4: Identification of HLA-associated naturally processed and presented epitopes

In chapter four I investigated whether epitopes from the C-peptide:WE14 polypeptide are naturally processed and presented by class II T1D-risk haplotypes. I did this by pulsing EBV B cells with pokeweed mitogen conjugated to the C-peptide:WE14 polypeptide (ADS established in chapter three) and used mass spectrometry to identify peptide epitopes.

Uptake and processing of exogenous proteins was evident, but more important was the identification of a pokeweed-derived peptide presented by HLA-DQ8. This finding indicates that the ADS enters the class II presentation pathway but is also a target for presenting naturally processed epitopes derived from the ADS system. However, I was unable to identify a hybrid epitope bound to HLA and consequently, in my experimental setting I do not have evidence that the polypeptide species is naturally processed and presented to generate a minimal hybrid epitope. Nonetheless, one might be able to use other antigen deliveries or other methods to evaluate whether this epitope is presented naturally for example, targeting of DEC205. As mentioned previously, delivery of a GAD65(274-286)-containing 41-mer by targeting DEC205 did not result in peptide presentation but there is the possibility that this delivery method may have been successful for the hybrid polypeptide. Theoretically, it would also be conceivable to deliver the C-peptide:WE14 ADS to EBV B cells and co-culture with a T cell clone which recognises the SLQPLAL-WSKMDQL epitope to further evaluate whether this epitope can be naturally processed from the polypeptide species. A T cell clone against this particular insulin-WE14 hybrid epitope is currently unavailable and ideally, the T cell clone

used in the aforementioned experiment would be isolated from the pancreatic islets of T1D patients as described by (Michels et al., 2017) and (Babon et al., 2016).

Currently, I conclude that the C-peptide:WE14 polypeptide is not naturally processed and presented but there are important caveats to contemplate. It might well be that it is processed and presented for example, taking into consideration that the amount of polypeptide present on the cell surface after pulsing may be lower than that of the ADS containing the GAD65(274-286) epitope (Figure 3.6B and 4.1A). In this way, the assay may be limiting the amount of peptide available for processing. Therefore, despite the efficiency of the ADS, from the start the odds were stacked against identifying an epitope derived from the polypeptide because there was less peptide starting material being delivered into the APC. To illustrate this further, it is important to highlight the five steps of antigen processing: capture, internalisation, processing, loading and export of peptide-HLA complexes to the cell surface (Neefjes et al., 2011). Using these five steps it is feasible to consider the differences between the GAD65(274-286)-containing ADS, which I have data to show is working well and the hybrid polypeptide ADS. Internalisation and export of peptide-HLA complexes to the cell surface are likely the same whereas differences may theoretically arise in the amount of peptide captured at the cell surface, loading of the processed peptides onto HLA molecules due to HLA binding affinities and processing of the two peptide species. Regarding the latter, I have generated data *in silico* to suggest that the C-peptide:WE14 polypeptide lacks candidate protease cleavage sites which may be important in deriving a hybrid epitope. Furthermore, given previous emphasis on the route of uptake one cannot exclude that a naturally processed hybrid epitope may be generated if the hybrid polypeptide was taken up into the APC via an alternative route.

Consistent with my data, evidence from the literature also suggests the C-peptide:WE14 HIP may not exist in the immunopeptidome. Despite being previously identified by (Delong et al., 2016) a recent publication by (Wan et al., 2020) could not identify the C-peptide:WE14 (LQTLAL-WSRMD) HIP in the cell surface MHC II immunopeptidome of pancreatic islets of NOD mice. Notwithstanding that β-cells from NOD mice may express MHC class II (IA-g7) (Walter et al., 2003; Zhao et al., 2015), (Wan et al., 2020) claim not that the islets present these epitopes but that the species that are eluted derive from MHC class II molecules present on APCs (macrophages and dendritic cells) within the islet. Although the authors are unable to identify the C-peptide:WE14 HIP and state that it may below the detection threshold, they are able to identify a C-peptide:IAPP2 HIP (LQTLAL-NAARD) in the MHC II peptidome. This finding would suggest that the C-peptide:WE14 HIP may be difficult to generate either because this particular epitope is not processed in the APC and/or it is not very prevalent in the crinosome (a candidate organelle for HIP formation) so is not available as a substrate. Indeed, a previous report did not identify the C-peptide:WE14 HIP species in the crinosome of pancreatic β-cells of NOD mice (Wan et al., 2018). If the C-peptide:WE14 HIP is not being generated in the islets its identification by (Delong et al., 2016) may be exclusive to the insulinoma cell line in which it was discovered. Nevertheless, the findings of (Wan et al., 2020) lend support to the notion identifying a naturally processed WE14-containing hybrid epitope bound to HLA may be challenging for a number of reasons. Most probably, a lack of or limited presentation of the C-peptide:WE14 epitope in APCs would also reduce the likelihood of a T cell response being elicited.

The identification of epitopes presented at the cell surface is to some extent determined by a hierarchy of available peptides. Evidence by (Wan et al., 2020) would suggest that hybrid peptides are present at very low abundance in the MHC II peptidome of islets and pancreatic

lymph nodes, representing less than 1% of the total peptidome in each case. This is in contrast to the initial findings of (Delong et al., 2016) in which the C-peptide:WE14 epitope is dominant in antigenic β-cell fractions, and it is suggested that molecular crowding within the β-cell might favour hybrid peptide generation. Knowing the former data, the C-peptide:WE14 peptide may not be ideal for use in the approaches I have been employing. The C-peptide:IAPP2 species however appears to be a strong HIP candidate as it has been identified both in the crinosome (Wan et al., 2018) and the class II immunopeptidome (Wan et al., 2020). It is possible that other HIPs are generated but because of their sequence motifs are rapidly cleaved/degraded, and hence are not very frequent in the class II immunopeptidome. The C-peptide:IAPP2 HIP may pose an exception, generating a hybrid intersection which is not targeted by other proteases. Similar experiments performed in chapter four could be employed for the C-peptide:IAPP2 HIP to investigate whether a naturally processed epitope can be generated from APCs. Whole blood assays using overlapping peptides and the generation of T cell lines/clones could be an alternative approach to identifying NPPEs. However, this merely shows whether T cells capable of responding to a specified peptide exist but does not provide proof that a peptide is in fact naturally processed and presented. This information can only be obtained by direct elution from class II HLA molecules.

6.4 Chapter 5: Peripheral blood reactivity to hybrid polypeptides

Using triple-colour FluoroSpot I set out to investigate whether immune responses against hybrid polypeptides can be identified in T1D patients. The questions addressed by comparing responses between the short epitope and the long hybrid polypeptide were firstly, whether the hybrid epitope is generated by processing of the hybrid polypeptide and secondly, whether it elicits a T cell response as a result. Because the TCR repertoire within a PBMC culture is diverse (Warren et al., 2011), some of which may be pre-activated, it is difficult to argue that

a response against the short and long version in the same individual is due to the same TCR specificity which recognises both. To do so would require a T cell clone of single specificity to assess whether it responds to the long and short versions in the same way.

I have generated data to show low-level reactivity against the C-peptide:WE14 polypeptide and in individuals which respond to the short hybrid epitope, not all of them elicit a response against the long polypeptide. This suggests that the short epitope is not easily generated from the long polypeptide. This finding was reproduced with another hybrid polypeptide, C-peptide:IAPP2 albeit responses were more frequent and it is possible that this may be attributable to differences in the binding capacity of WE14 versus IAPP2. WE14 does not bind well into the MHC class II binding groove (Stadinski et al., 2010a) whereas peptides derived from IAPP might bind better (Baker et al., 2013). In this case it is important to refer back to the five steps of antigen processing; these data indicate uptake, processing or presentation of long hybrid polypeptides is limited and/or inefficient and does not lead to the same peptide binding as is the case for the shortened epitope. As uptake represents a key step, the route by which APCs acquire the polypeptide in the FluoroSpot system is likely the more physiologically relevant route compared with using an ADS. However, few patients elicited a positive cytokine response to the C-peptide:WE14 polypeptide and equally in the ADS system, which utilised a different route of uptake, a C-peptide:WE14 hybrid epitope was not eluted.

Evidence exists suggesting that long peptides can potentially bypass the need for processing. Namely, (Wan et al., 2020) showed that 29-mer peptides can be eluted from APCs and do not require intracellular processing. While the hybrid polypeptides used in my FluoroSpot assays are longer than 29 amino acid residues, potentially an argument could be made that they bind the cell surface without needing processing. Although there is evidence to suggest that peptides

do not have to go through the internalisation processing pathway, data does not exist to exclude that this is still a crucial pathway operative within the assay. Indeed, the C-peptide:WE14 polypeptide species needs processing to elicit a T cell response. This can be reasoned because conjugation of pokeweed mitogen to C-peptide:WE14 in the ADS is not 100% efficient, thus I would expect free unconjugated peptide to be present during pulsing. If unconjugated peptide was in fact available for direct binding to HLA, it is technically feasible that I would be able to detect these peptides (as the LC-MS/MS technique used detects peptides up to 45 amino acids long and the test polypeptide was 41 amino acids).

I also found that the orientation of the hybrid polypeptide influences the immune response. Further, I was interested in understanding why changes in the orientation of the polypeptide lead to a greater prevalence of responses. Hybrid polypeptides comprise either the normal orientation (C-peptide:WE14) or the reversed orientation (WE14:C-peptide). Using a predictive algorithm, I generate data to show that the binding cores of reversed orientation polypeptides contain the intersection of the two granule peptides. Under natural circumstances, a natural cleavage product of ChgA is the WE14 sequence (WSKMDQLAKELTAE) (Orr et al., 2002). Consequently, the N-terminal region of this species is inevitably present and providing C-peptide is cleaved at the C-terminus, a reaction between the two species could potentially occur and a N-terminal C-peptide (SLQPLAL-WSKMDQL) epitope is generated. To generate an epitope in the reverse orientation (WSKMDQL-SLQPLAL) ChgA would need to be cleaved not at the natural cleavage site, but internally within the WE14 sequence and likewise the case for C-peptide. Cleavage within the WE14 and C-peptide sequence may not occur at the same rate as the conventional cleavage site (Muller et al., 2009). Therefore, the availability of both peptides at the same time for a reaction to occur would be lower. Unless such epitopes are generated at a later stage during degradation, it is then less likely that the

reversed orientation epitope is presented leading to reduced opportunity for tolerance induction. To further establish the region of the reversed hybrid polypeptide that is targeted, it would be necessary to also test the short-reversed hybrid epitopes by for example, IFN-γ ELISpot.

I have some evidence from my FluoroSpot data that the C-peptide:WE14 polypeptide is capable of eliciting a T cell response which suggests there is some degree of *in vivo* processing. Nevertheless, I have been unable to show the C-peptide:WE14 polypeptide is processed naturally in my setting. I generate data to indicate that T cells which recognise the C-peptide:WE14 polypeptide are present in a few individuals, but it is possible that those T cells could also be specific for other unrelated targets, as a result of TCR promiscuity (Wooldridge et al., 2012). Sequences of both the normal and reversed orientation hybrid polypeptides could mimic the sequence of a pathogen, for example. However, no significant similarly to amino acid sequences present in the proteome of bacteria, archaea and viruses was identified using the National Centre for Biotechnology Information (NCBI) BLAST server and the Virus Pathogen Database and Analysis Resource (ViPR) (Pickett et al., 2012).

Non-disease controls were not included in this study and therefore I do not have evidence to suggest the observed response is specific to disease. Accordingly, the data presented does not support the notion that the C-peptide:WE14 hybrid polypeptide is processed and presented and might therefore be disease-relevant. If one wanted to investigate whether responses to hybrid (poly)peptides are disease-relevant, a larger cohort of patients and matched controls would be required but this was beyond the scope of the project. Assuming 0% of responders in controls and 20% in patients (based on (Baker et al., 2019b); 2/11 respond to hEL:ChgA-WE14 [SLQPLAL-WSKMDQL] epitope) a sample size of 32 in each group would be necessary to

detect significant differences in the prevalence of responses with 80% power. Furthermore, from my data I cannot make conclusions regarding whether responses to hybrid (poly)peptides are more prevalent and/or of greater magnitude in T1D patients compared to other naturally processed epitopes restricted by the same HLA class II molecule.

6.5 Implications of thesis findings on the disease model

Overall the data would suggest that hybrid epitopes are not generated through conventional processing and presentation of a long hybrid polypeptide. It is possible that shorter fragments are being generated and then they are being fused, and this is likely happening within β-cells. Indeed, a model of unconventional epitope generation within the β-cell has been proposed based on the observation that the β-cell has the capacity to secrete peptides into the periphery which may or may not require further processing (Wan et al., 2018). The subcellular location within the β-cell where hybrid peptides are generated is still not clear but granular compartments such as the lysosome (Wang et al., 2019), crinosome (Wan et al., 2018) or insulin granule (Delong et al., 2016) are candidate organelles. Besides the proteasome, other proteases such as asparaginyl endoprotease (AEP) are known to catalyse transpeptidation, and in particular, the human homolog of AEP which is also involved in the generation of class II-restricted epitopes (Manoury et al., 1998). Normal or stress-induced turnover of granules by crinophagy (Sandberg and Borg, 2006; Sobota et al., 2009) represents an attractive conduit to hybrid peptide formation as such crinosomes contain proteolytic enzymes derived from both lysosomes and aged insulin granules. How efficiently such proteases perform this reaction relative to the proteasome, which has several characteristics that favour catalysis of peptide splicing (Liepe et al., 2010) is not understood.

The question then remains how hybrid peptides are delivered into the class II pathway. Low binding affinity of CLIP to IA-g7, owing to the non–aspartic acid residue at position 57, permits spontaneous replacement of CLIP at the cell surface by extracellular peptides (Ito et al., 2018). In the context of HLA-DQ2 and -DQ8, both of which share the β57 polymorphism in the binding pocket with IA-g7 (Todd et al., 1987), low affinity binding of CLIP to these molecules has also been shown (Fallang et al., 2008; Reed et al., 1997; Wiesner et al., 2008). For this reason, extracellular replacement of CLIP bound to HLA-DQ2/8 molecules may also be operative. Together with the findings of (Wan et al., 2020) whereby use of an endocytic/processing inhibitor still lead to loading of long MHC II peptides on the cell surface, these studies suggest that there is pathway that allows for epitopes to access HLA class II molecules without the need for internalisation and processing.

Using this knowledge, the model emerging is that within the β-cell processing is occurring, pre-processed peptides are secreted and can be directly loaded onto HLA molecules on the surface of APCs e.g. DCs (Figure 6.1A, B). Evidence in the NOD mouse suggests the concentration of epitopes within the peripheral lymph nodes draining the islet, is high enough to be presented by APCs and interact with T cells (Wan et al., 2018). Currently it is unclear whether this is also true for humans, but certainly within the lymph node itself close to the islet there is a high concentration of secreted peptides which may promote extracellular peptide displacement at the cell surface. Furthermore, the model implies that for cell surface displacement of HLA-DQ2 and -DQ8 molecules, extracellular peptides would need to possess a high binding affinity (Delong et al., 2016; Ito et al., 2018). This would suggest a selection process exists whereby only peptides with medium-to-high affinity will gain access to the HLA pathway maybe not through internalisation.

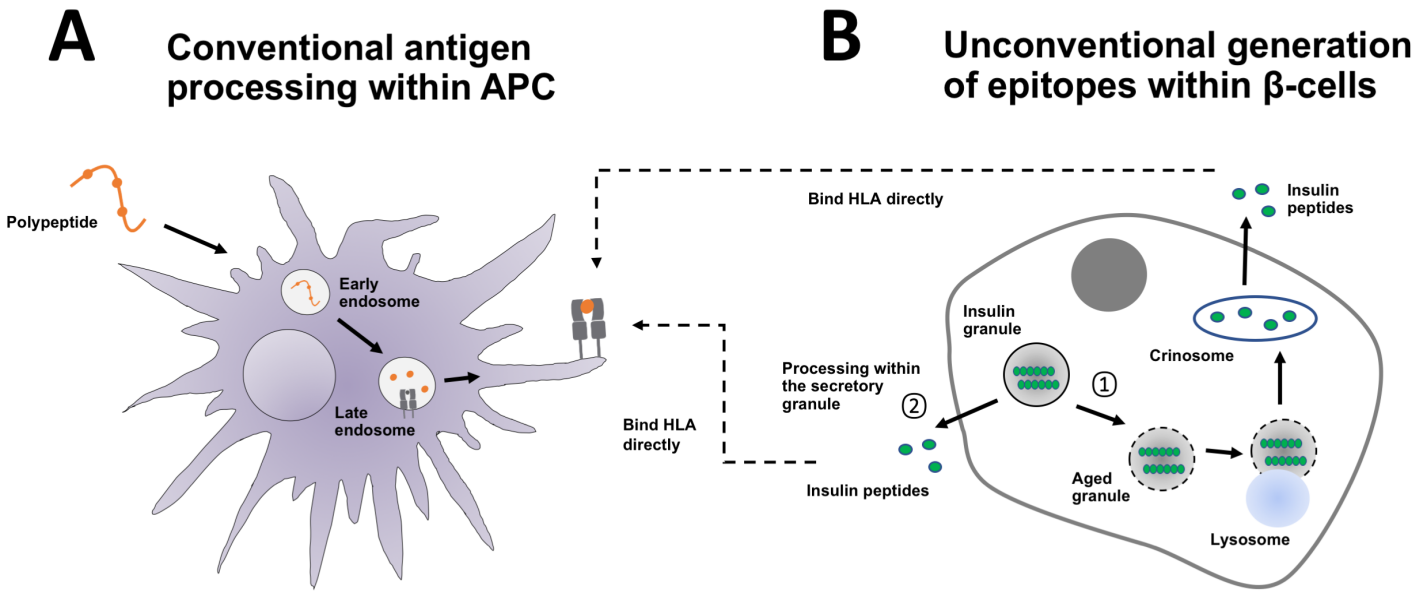


Figure 6.1 Different modes of antigen processing. Within an APC a polypeptide is taken up into endosomal compartments, processed into peptides and loaded onto HLA molecules (A). Peptide-HLA complexes are subsequently transported to the cell surface. The β-cell can process insulin peptides within the crinosome (①) formed by the fusion of lysosomes and aged insulin granules, but also within the secretory granule (②) (observation in Peakman lab). In either event, these peptides are released and can bind HLA molecules by extracellular displacement (B). APC, antigen-presenting cell; HLA, human leukocyte antigen.

An attractive concept of the model is that there is disease relevance in HLA-DQ2 and -DQ8 over other class II molecules. Specifically, the mode of antigen acquisition to HLA might distinguish where these hybrid peptides are more relevant. Although evidence exists that HIPs can be HLA-DR-restricted (Baker et al., 2019b), rapid CLIP dissociation has only been reported to occur for IA-g7 due to the β57 polymorphism (Ito et al., 2018) which is structurally similar to HLA-DQ8 and -DQ2 (Lee et al., 2001). Therefore, the acquisition of hybrid peptides into HLA-DR, or any other HLA class II molecule for that matter, is probably restricted by the notion that CLIP binds relatively strong (Amicosante et al., 2009; Hotzel and Koch, 1996) making peptide exchange difficult to accomplish. Thus, HLA-DQ is likely the most relevant HLA class II molecule to acquire hybrid peptides in the unconventional generation pathway.

6.6 Future outlook

This thesis aimed to understand how a hybrid epitope can be derived and points toward an unconventional mechanism of generation. Immune responses against hybrid peptides have been described in the peripheral blood (Baker et al., 2019b), but in this thesis I do not have evidence to suggest that they are naturally processed and presented from a long polypeptide in an antigen-presenting cell. Therefore, the β-cell is the most likely origin of hybrid peptides and future discovery of hybrid epitopes should focus on the peptide repertoire on APCs under conditions where they have the opportunity to take up those hybrid peptides. Chapter four would suggest that further studies should utilise broader and more complex methodologies for example, different routes of antigen delivery into the APC and/or other APCs aside from B cells, to interrogate the HIP landscape more deeply. The jury is still out as to whether hybrid peptides are key drivers or initiators of disease. At least for the hybrid peptides identified so far, I would not predict that responses to hybrid peptides are a key driver of disease since they do not appear to elicit a higher prevalence or magnitude of response than that of other disease-relevant epitopes (Peakman, unpublished). Studying at-risk individuals before any overt signs of autoimmunity and following its progression in parallel with measurement of immune responses to native and hybrid peptides may prove useful in the future to discern pathological pathways.

REFERENCES

Abelin, J.G., D. Harjanto, M. Malloy, P. Suri, T. Colson, S.P. Goulding, A.L. Creech, L.R. Serrano, G. Nasir, Y. Nasrullah, C.D. McGann, D. Velez, Y.S. Ting, A. Poran, D.A. Rothenberg, S. Chhangawala, A. Rubinsteyn, J. Hammerbacher, R.B. Gaynor, E.F. Fritsch, J. Greshock, R.C. Oslund, D. Barthelme, T.A. Addona, C.M. Arieta, and M.S. Rooney. 2019. Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. *Immunity* 51:766-779 e717.

Achenbach, P., V. Lampasona, U. Landherr, K. Koczwara, S. Krause, H. Grallert, C. Winkler, M. Pfluger, T. Illig, E. Bonifacio, and A.G. Ziegler. 2009. Autoantibodies to zinc transporter 8 and SLC30A8 genotype stratify type 1 diabetes risk. *Diabetologia* 52:1881-1888.

Alhadj Ali, M., Y.F. Liu, S. Arif, D. Tatovic, H. Shariff, V.B. Gibson, N. Yusuf, R. Baptista, M. Eichmann, N. Petrov, S. Heck, J.H.M. Yang, T.I.M. Tree, I. Pujol-Autonell, L. Yeo, L.R. Baumard, R. Stenson, A. Howell, A. Clark, Z. Boult, J. Powrie, L. Adams, F.S. Wong, S. Luzio, G. Dunseath, K. Green, A. O'Keefe, G. Bayly, N. Thorogood, R. Andrews, N. Leech, F. Joseph, S. Nair, S. Seal, H. Cheung, C. Beam, R. Hills, M. Peakman, and C.M. Dayan. 2017. Metabolic and immune effects of immunotherapy with proinsulin peptide in human new-onset type 1 diabetes. *Sci Transl Med* 9:

American Diabetes, A. 2018. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018. *Diabetes Care* 41:S13-S27.

Amicosante, M., F. Berretta, R. Dweik, and C. Saltini. 2009. Role of high-affinity HLA-DP specific CLIP-derived peptides in beryllium binding to the HLA-DPGlu69 berylliosis-associated molecules and presentation to beryllium-sensitized T cells. *Immunology* 128:e462-470.

Anderson, A.C., L.B. Nicholson, K.L. Legge, V. Turchin, H. Zaghouani, and V.K. Kuchroo. 2000. High frequency of autoreactive myelin proteolipid protein-specific T cells in the periphery of naive mice: mechanisms of selection of the self-reactive repertoire. *J Exp Med* 191:761-770.

Anderson, M.S., E.S. Venanzi, L. Klein, Z. Chen, S.P. Berzins, S.J. Turley, H. von Boehmer, R. Bronson, A. Dierich, C. Benoist, and D. Mathis. 2002. Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298:1395-1401.

Anton, L.C., and J.W. Yewdell. 2014. Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors. *J Leukoc Biol* 95:551-562.

Arentz-Hansen, H., R. Körner, Ø. Molberg, H. Quarsten, W. Vader, Y.M.C. Kooy, K.E.A. Lundin, F. Koning, P. Roepstorff, L.M. Sollid, and S.N. McAdam. 2000. The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *The Journal of Experimental Medicine* 191:603-612.

Arif, S., T.I. Tree, T.P. Astill, J.M. Tremble, A.J. Bishop, C.M. Dayan, B.O. Roep, and M. Peakman. 2004. Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health. *J Clin Invest* 113:451-463.

Babiceanu, M., F. Qin, Z. Xie, Y. Jia, K. Lopez, N. Janus, L. Facemire, S. Kumar, Y. Pang, Y. Qi, I.M. Lazar, and H. Li. 2016. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res* 44:2859-2872.

Babon, J.A., M.E. DeNicola, D.M. Blodgett, I. Crevecoeur, T.S. Buttrick, R. Maehr, R. Bottino, A. Naji, J. Kaddis, W. Elyaman, E.A. James, R. Haliyur, M. Brissova, L. Overbergh, C. Mathieu, T. Delong, K. Haskins, A. Pugliese, M. Campbell-Thompson, C. Mathews, M.A. Atkinson, A.C. Powers, D.M. Harlan, and S.C. Kent. 2016. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. *Nat Med* 22:1482-1487.

Baker, R.L., T. Delong, G. Barbour, B. Bradley, M. Nakayama, and K. Haskins. 2013. Cutting edge: CD4 T cells reactive to an islet amyloid polypeptide peptide accumulate in the pancreas and contribute to disease pathogenesis in nonobese diabetic mice. *J Immunol* 191:3990-3994.

Baker, R.L., B.L. Jamison, and K. Haskins. 2019a. Hybrid insulin peptides are neo-epitopes for CD4 T cells in autoimmune diabetes. *Curr Opin Endocrinol Diabetes Obes* 26:195-200.

Baker, R.L., B.L. Jamison, T.A. Wiles, R.S. Lindsay, G. Barbour, B. Bradley, T. Delong, R.S. Friedman, M. Nakayama, and K. Haskins. 2018. CD4 T Cells Reactive to Hybrid Insulin Peptides Are Indicators of Disease Activity in the NOD Mouse. *Diabetes* 67:1836-1846.

Baker, R.L., M. Rihanek, A.C. Hohenstein, M. Nakayama, A. Michels, P.A. Gottlieb, K. Haskins, and T. Delong. 2019b. Hybrid Insulin Peptides Are Autoantigens in Type 1 Diabetes. *Diabetes* 68:1830-1840.

Belizaire, R., and E.R. Unanue. 2009. Targeting proteins to distinct subcellular compartments reveals unique requirements for MHC class I and II presentation. *Proc Natl Acad Sci U S A* 106:17463-17468.

Bergseng, E., S. Dorum, M.O. Arntzen, M. Nielsen, S. Nygard, S. Buus, G.A. de Souza, and L.M. Sollid. 2015. Different binding motifs of the celiac disease-associated HLA molecules DQ2.5, DQ2.2, and DQ7.5 revealed by relative quantitative proteomics of endogenous peptide repertoires. *Immunogenetics* 67:73-84.

Berkers, C.R., A. de Jong, H. Ovaa, and B. Rodenko. 2009. Transpeptidation and reverse proteolysis and their consequences for immunity. *Int J Biochem Cell Biol* 41:66-71.

Berkers, C.R., A. de Jong, K.G. Schuurman, C. Linnemann, H.D. Meiring, L. Janssen, J.J. Neefjes, T.N. Schumacher, B. Rodenko, and H. Ovaa. 2015. Definition of Proteasomal Peptide Splicing Rules for High-Efficiency Spliced Peptide Presentation by MHC Class I Molecules. *J Immunol* 195:4085-4095.

Bhattarai, S., I.F. Godsland, S. Misra, D.G. Johnston, and N. Oliver. 2019. Metabolic health and vascular complications in type 1 diabetes. *J Diabetes Complications* 33:634-640.

Blum, J.S., P.A. Wearsch, and P. Cresswell. 2013. Pathways of antigen processing. *Annu Rev Immunol* 31:443-473.

Bonasio, R., M.L. Scimone, P. Schaerli, N. Grabie, A.H. Lichtman, and U.H. von Andrian. 2006. Clonal deletion of thymocytes by circulating dendritic cells homing to the thymus. *Nat Immunol* 7:1092-1100.

Bondinas, G.P., A.K. Moustakas, and G.K. Papadopoulos. 2007. The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. *Immunogenetics* 59:539-553.

Borbulevych, O.Y., T.K. Baxter, Z. Yu, N.P. Restifo, and B.M. Baker. 2005. Increased immunogenicity of an anchor-modified tumor-associated antigen is due to the enhanced stability of the peptide/MHC complex: implications for vaccine design. *J Immunol* 174:4812-4820.

Bradford, C.M., I. Ramos, A.K. Cross, G. Haddock, S. McQuaid, A.P. Nicholas, and M.N. Woodroofe. 2014. Localisation of citrullinated proteins in normal appearing white matter and lesions in the central nervous system in multiple sclerosis. *J Neuroimmunol* 273:85-95.

Brunner, Y., Y. Coute, M. Iezzi, M. Foti, M. Fukuda, D.F. Hochstrasser, C.B. Wollheim, and J.C. Sanchez. 2007. Proteomics analysis of insulin secretory granules. *Mol Cell Proteomics* 6:1007-1017.

Bulek, A.M., D.K. Cole, A. Skowera, G. Dolton, S. Gras, F. Madura, A. Fuller, J.J. Miles, E. Gostick, D.A. Price, J.W. Drijfhout, R.R. Knight, G.C. Huang, N. Lissin, P.E. Molloy, L. Wooldridge, B.K. Jakobsen, J. Rossjohn, M. Peakman, P.J. Rizkallah, and A.K. Sewell. 2012. Structural basis for the killing of human beta cells by CD8(+) T cells in type 1 diabetes. *Nat Immunol* 13:283-289.

Burrack, A.L., T. Martinov, and B.T. Fife. 2017. T Cell-Mediated Beta Cell Destruction: Autoimmunity and Alloimmunity in the Context of Type 1 Diabetes. *Front Endocrinol (Lausanne)* 8:343.

Calender, A., M. Billaud, J.P. Aubry, J. Banchereau, M. Vuillaume, and G.M. Lenoir. 1987. Epstein-Barr virus (EBV) induces expression of B-cell activation markers on in vitro infection of EBV-negative B-lymphoma cells. *Proceedings of the National Academy of Sciences* 84:8060-8064.

Campbell, P.J., P.J. Stephens, E.D. Pleasance, S. O'Meara, H. Li, T. Santarius, L.A. Stebbings, C. Leroy, S. Edkins, C. Hardy, J.W. Teague, A. Menzies, I. Goodhead, D.J. Turner, C.M. Clee, M.A. Quail, A. Cox, C. Brown, R. Durbin, M.E. Hurles, P.A. Edwards, G.R. Bignell, M.R. Stratton, and P.A. Futreal. 2008. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet* 40:722-729.

Carrero, J.A., D.P. McCarthy, S.T. Ferris, X. Wan, H. Hu, B.H. Zinselmeyer, A.N. Vomund, and E.R. Unanue. 2017. Resident macrophages of pancreatic islets have a seminal role in the initiation of autoimmune diabetes of NOD mice. *Proc Natl Acad Sci U S A* 114:E10418-E10427.

Chiang, J.L., M.S. Kirkman, L.M. Laffel, A.L. Peters, and A. Type 1 Diabetes Sourcebook. 2014. Type 1 diabetes through the life span: a position statement of the American Diabetes Association. *Diabetes Care* 37:2034-2054.

Chicz, R.M., R.G. Urban, J.C. Gorga, D.A. Vignali, W.S. Lane, and J.L. Strominger. 1993. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med* 178:27-47.

Chicz, R.M., R.G. Urban, W.S. Lane, J.C. Gorga, L.J. Stern, D.A. Vignali, and J.L. Strominger. 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* 358:764-768.

Chilson, O.P., and A.E. Kelly-Chilson. 1989. Mitogenic lectins bind to the antigen receptor on human lymphocytes. *Eur J Immunol* 19:389-396.

Cho, K.J., and P.A. Roche. 2013. Regulation of MHC Class II-Peptide Complex Expression by Ubiquitination. *Front Immunol* 4:369.

Clements, G.B., D.N. Galbraith, and K.W. Taylor. 1995. Coxsackie B virus infection and onset of childhood diabetes. *Lancet* 346:221-223.

Cohn, L., B. Chatterjee, F. Esselborn, A. Smed-Sorensen, N. Nakamura, C. Chalouni, B.C. Lee, R. Vandlen, T. Keler, P. Lauer, D. Brockstedt, I. Mellman, and L. Delamarre. 2013. Antigen delivery to early endosomes eliminates the superiority of human blood BDCA3+ dendritic cells at cross presentation. *J Exp Med* 210:1049-1063.

Colaert, N., K. Helsens, L. Martens, J. Vandekerckhove, and K. Gevaert. 2009. Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 6:786-787.

Cole, D.K., E.S. Edwards, K.K. Wynn, M. Clement, J.J. Miles, K. Ladell, J. Ekeruche, E. Gostick, K.J. Adams, A. Skowera, M. Peakman, L. Wooldridge, D.A. Price, and A.K. Sewell. 2010. Modification of MHC anchor residues generates heteroclitic peptides that alter TCR binding and T cell recognition. *J Immunol* 185:2600-2610.

Couper, J., and K.C. Donaghue. 2009. Phases of diabetes in children and adolescents. *Pediatr Diabetes* 10 Suppl 12:13-16.

Craig, M.E., A. Hattersley, and K.C. Donaghue. 2009. Definition, epidemiology and classification of diabetes in children and adolescents. *Pediatr Diabetes* 10 Suppl 12:3-12.

Crawford, F., B. Stadinski, N. Jin, A. Michels, M. Nakayama, P. Pratt, P. Marrack, G. Eisenbarth, and J.W. Kappler. 2011. Specificity and detection of insulin-reactive CD4+ T cells in type 1 diabetes in the nonobese diabetic (NOD) mouse. *Proc Natl Acad Sci U S A* 108:16729-16734.

Cresswell, P. 1996. Invariant Chain Structure and MHC Class II Function. *Cell* 84:505-507.

Croft, N.P., S.A. Smith, Y.C. Wong, C.T. Tan, N.L. Dudek, I.E. Flesch, L.C. Lin, D.C. Tscharke, and A.W. Purcell. 2013. Kinetics of antigen expression and epitope presentation during virus infection. *PLoS Pathog* 9:e1003129.

Cusick, M.F., J.E. Libbey, and R.S. Fujinami. 2012. Molecular mimicry as a mechanism of autoimmune disease. *Clin Rev Allergy Immunol* 42:102-111.

Dalet, A., V. Stroobant, N. Vigneron, and B.J. Van den Eynde. 2011. Differences in the production of spliced antigenic peptides by the standard proteasome and the immunoproteasome. *Eur J Immunol* 41:39-46.

Dalet, A., N. Vigneron, V. Stroobant, K. Hanada, and B.J. Van den Eynde. 2010. Splicing of distant peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J Immunol* 184:3016-3024.

Delong, T., R.L. Baker, J. He, G. Barbour, B. Bradley, and K. Haskins. 2012. Diabetogenic T-cell clones recognize an altered peptide of chromogranin A. *Diabetes* 61:3239-3246.

Delong, T., T.A. Wiles, R.L. Baker, B. Bradley, G. Barbour, R. Reisdorph, M. Armstrong, R.L. Powell, N. Reisdorph, N. Kumar, C.M. Elso, M. DeNicola, R. Bottino, A.C. Powers, D.M. Harlan, S.C. Kent, S.I. Mannering, and K. Haskins. 2016. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* 351:711-714.

Denroche, H.C., and C.B. Verchere. 2018. IAPP and type 1 diabetes: implications for immunity, metabolism and islet transplants. *J Mol Endocrinol* 60:R57-R75.

Denzin, L.K., J.L. Fallas, M. Prendes, and W. Yi. 2005. Right place, right time, right peptide: DO keeps DM focused. *Immunol Rev* 207:279-292.

Diamond Project Group. 2006. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabet Med* 23:857-866.

Diana, J., Y. Simoni, L. Furio, L. Beaudoin, B. Agerberth, F. Barrat, and A. Lehuen. 2013. Crosstalk between neutrophils, B-1a cells and plasmacytoid dendritic cells initiates autoimmune diabetes. *Nat Med* 19:65-73.

Dianzani, C., E. Bellavista, J. Liepe, C. Verderio, M. Martucci, A. Santoro, A. Chiocchetti, C.L. Gigliotti, E. Boggio, B. Ferrara, L. Riganti, C. Keller, K. Janek, A. Niewienda, C. Fenoglio, M. Sorosina, R. Cantello, P.M. Kloetzel, M.P. Stumpf, F. Paul, K. Ruprecht, D. Galimberti, F. Martinelli Boneschi, C. Comi, U. Dianzani, and M. Mishto. 2017. Extracellular proteasome-osteopontin circuit regulates cell migration with implications in multiple sclerosis. *Sci Rep* 7:43718.

Dunne, J.L., S.J. Richardson, M.A. Atkinson, M.E. Craig, K. Dahl-Jorgensen, M. Flodstrom-Tullberg, H. Hyoty, R.A. Insel, A. Lernmark, R.E. Lloyd, N.G. Morgan, and A. Pugliese. 2019. Rationale for enteroviral vaccination and antiviral therapies in human type 1 diabetes. *Diabetologia* 62:744-753.

Dwivedi, N., I. Neeli, N. Schall, H. Wan, D.M. Desiderio, E. Csernok, P.R. Thompson, H. Dali, J.P. Briand, S. Muller, and M. Radic. 2014. Deimination of linker histones links neutrophil extracellular trap release with autoantibodies in systemic autoimmunity. *FASEB J* 28:2840-2851.

Ebstein, F., K. Textoris-Taube, C. Keller, R. Golnik, N. Vigneron, B.J. Van den Eynde, B. Schuler-Thurner, D. Schadendorf, F.K. Lorenz, W. Uckert, S. Urban, A. Lehmann, N. Albrecht-Koepke, K. Janek, P. Henklein, A. Niewienda, P.M. Kloetzel, and M. Mishto. 2016. Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. *Sci Rep* 6:24032.

Erlich, H., A.M. Valdes, J. Noble, J.A. Carlson, M. Varney, P. Concannon, J.C. Mychaleckyj, J.A. Todd, P. Bonella, A.L. Fear, E. Lavant, A. Louey, P. Moonsamy, and C. Type 1 Diabetes Genetics. 2008. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. *Diabetes* 57:1084-1092.

EURODIAB ACE Study Group. 2000. Variation and trends in incidence of childhood diabetes in Europe. *The Lancet* 355:873-876.

Fallang, L.E., S. Roh, A. Holm, E. Bergseng, T. Yoon, B. Fleckenstein, A. Bandyopadhyay, E.D. Mellins, and L.M. Sollid. 2008. Complexes of two cohorts of CLIP peptides and HLA-DQ2 of the autoimmune DR3-DQ2 haplotype are poor substrates for HLA-DM. *J Immunol* 181:5451-5461.

Faridi, P., C. Li, S.H. Ramarathinam, J.P. Vivian, P.T. Illing, N.A. Mifsud, R. Ayala, J. Song, L.J. Gearing, P.J. Hertzog, N. Ternette, J. Rossjohn, N.P. Croft, and A.W. Purcell. 2018. A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. *Sci Immunol* 3:

Ferris, S.T., J.A. Carrero, J.F. Mohan, B. Calderon, K.M. Murphy, and E.R. Unanue. 2014. A minor subset of Batf3-dependent antigen-presenting cells in islets of Langerhans is essential for the development of autoimmune diabetes. *Immunity* 41:657-669.

Figdor, C.G., Y. van Kooyk, and G.J. Adema. 2002. C-type lectin receptors on dendritic cells and Langerhans cells. *Nat Rev Immunol* 2:77-84.

Friman, G., J. Fohlman, G. Frisk, H. Diderholm, U. Ewald, M. Kobbah, and T. Tuvemo. 1985. An incidence peak of juvenile diabetes. Relation to Coxsackie B virus immune response. *Acta Paediatr Scand Suppl* 320:14-19.

Garcia-Vallejo, J.J., K. Bloem, L.M. Knippels, J. Garssen, S.J. van Vliet, and Y. van Kooyk. 2015. The Consequences of Multiple Simultaneous C-Type Lectin-Ligand Interactions: DCIR Alters the Endo-Lysosomal Routing of DC-SIGN. *Front Immunol* 6:87.

Garg, G., J.R. Tyler, J.H. Yang, A.J. Cutler, K. Downes, M. Pekalski, G.L. Bell, S. Nutland, M. Peakman, J.A. Todd, L.S. Wicker, and T.I. Tree. 2012. Type 1 diabetes-associated IL2RA variation lowers IL-2 signaling and contributes to diminished CD4+CD25+ regulatory T cell function. *J Immunol* 188:4644-4653.

Godkin, A., T. Friede, M. Davenport, S. Stevanovic, A. Willis, D. Jewell, A. Hill, and H.G. Rammensee. 1997. Use of eluted peptide sequence data to identify the binding characteristics of peptides to the insulin-dependent diabetes susceptibility allele HLA-DQ8 (DQ 3.2). *Int Immunol* 9:905-911.

Gonzalez-Duque, S., M.E. Azoury, M.L. Colli, G. Afonso, J.V. Turatsinze, L. Nigi, A.I. Lalanne, G. Sebastiani, A. Carre, S. Pinto, S. Culina, N. Corcos, M. Bugliani, P. Marchetti, M. Armanet, M. Diedisheim, B. Kyewski, L.M. Steinmetz, S. Buus, S. You, D. Dubois-Laforgue, E. Larger, J.P. Beressi, G. Bruno, F. Dotta, R. Scharfmann, D.L. Eizirik, Y. Verdier, J. Vinh, and R. Mallone. 2018. Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. *Cell Metab* 28:946-960 e946.

Gordon, S. 2016. Phagocytosis: An Immunobiologic Process. *Immunity* 44:463-475.

Gottlieb, P.A., T. Delong, R.L. Baker, L. Fitzgerald-Miller, R. Wagner, G. Cook, M.R. Rewers, A. Michels, and K. Haskins. 2014. Chromogranin A is a T cell antigen in human type 1 diabetes. *J Autoimmun* 50:38-41.

Grant, B.D., and J.G. Donaldson. 2009. Pathways and mechanisms of endocytic recycling. *Nat Rev Mol Cell Biol* 10:597-608.

Green, A., E.A. Gale, and C.C. Patterson. 1992. Incidence of childhood-onset insulin-dependent diabetes mellitus: the EURODIAB ACE Study. *Lancet* 339:905-909.

Greer, J.M., B. Denis, R.A. Sobel, and E. Trifilieff. 2001. Thiopalmitoylation of myelin proteolipid protein epitopes enhances immunogenicity and encephalitogenicity. *J Immunol* 166:6907-6913.

Groettrup, M., C.J. Kirk, and M. Basler. 2010. Proteasomes in immune cells: more than peptide producers? *Nat Rev Immunol* 10:73-78.

Gubin, M.M., X. Zhang, H. Schuster, E. Caron, J.P. Ward, T. Noguchi, Y. Ivanova, J. Hundal, C.D. Arthur, W.J. Krebber, G.E. Mulder, M. Toebes, M.D. Vesely, S.S. Lam, A.J. Korman, J.P. Allison, G.J. Freeman, A.H. Sharpe, E.L. Pearce, T.N. Schumacher, R. Aebersold, H.G. Rammensee, C.J. Melief, E.R. Mardis, W.E. Gillanders, M.N. Artyomov, and R.D. Schreiber. 2014. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 515:577-581.

Guéry, J., and L. Adorini. 1995. Dendritic cells are the most efficient in presenting endogenous naturally processed self-epitopes to class II-restricted T cells. *J Immunol* 154:536-544.

Gunasekaran, U., C.W. Hudgens, B.T. Wright, M.F. Maulis, and M. Gannon. 2012. Differential regulation of embryonic and adult beta cell replication. *Cell Cycle* 11:2431-2442.

Gurer, C., T. Strowig, F. Brilot, M. Pack, C. Trumpfheller, F. Arrey, C.G. Park, R.M. Steinman, and C. Munz. 2008. Targeting the nuclear antigen 1 of Epstein-Barr virus to the human endocytic receptor DEC-205 stimulates protective T-cell responses. *Blood* 112:1231-1239.

Hadeiba, H., K. Lahl, A. Edalati, C. Oderup, A. Habtezion, R. Pachynski, L. Nguyen, A. Ghodsi, S. Adler, and E.C. Butcher. 2012. Plasmacytoid dendritic cells transport peripheral antigens to the thymus to promote central tolerance. *Immunity* 36:438-450.

Hanada, K., J.W. Yewdell, and J.C. Yang. 2004. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* 427:252-256.

Harbige, J., M. Eichmann, and M. Peakman. 2017. New insights into non-conventional epitopes as T cell targets: The missing link for breaking immune tolerance in autoimmune disease? *J Autoimmun* 84:12-20.

Hickey, A.J., J.W. Bradley, G.L. Skea, M.J. Middleditch, C.M. Buchanan, A.R. Phillips, and G.J. Cooper. 2009. Proteins associated with immunopurified granules from a model pancreatic islet beta-cell system: proteomic snapshot of an endocrine secretory granule. *J Proteome Res* 8:178-186.

Hogquist, K.A., T.A. Baldwin, and S.C. Jameson. 2005. Central tolerance: learning self-control in the thymus. *Nat Rev Immunol* 5:772-782.

Holland, C.J., G. Dolton, M. Scurr, K. Ladell, A.J. Schauenburg, K. Miners, F. Madura, A.K. Sewell, D.A. Price, D.K. Cole, and A.J. Godkin. 2015. Enhanced Detection of Antigen-Specific CD4+ T Cells Using Altered Peptide Flanking Residue Peptide-MHC Class II Multimers. *J Immunol* 195:5827-5836.

Holohan, D.R., F. Van Gool, and J.A. Bluestone. 2019. Thymically-derived Foxp3+ regulatory T cells are the primary regulators of type 1 diabetes in the non-obese diabetic mouse model. *PLoS One* 14:e0217728.

Hotzel, C., and N. Koch. 1996. The invariant chain derived fragment CLIP is an efficient in vitro inhibitor of peptide binding to MHC class II molecules. *Molecular Immunology* 33:25-31.

Hou, P., E. Araujo, T. Zhao, M. Zhang, D. Massenburg, M. Veselits, C. Doyle, A.R. Dinner, and M.R. Clark. 2006. B cell antigen receptor signaling and internalization are mutually exclusive events. *PLoS Biol* 4:e200.

Huang, J., Y. Xiao, A. Xu, and Z. Zhou. 2016. Neutrophils in type 1 diabetes. *J Diabetes Investig* 7:652-663.

Inaba, H., W. Martin, M. Ardito, A.S. De Groot, and L.J. De Groot. 2010. The role of glutamic or aspartic acid in position four of the epitope binding motif and thyrotropin receptor-extracellular domain epitope selection in Graves' disease. *J Clin Endocrinol Metab* 95:2909-2916.

Ito, Y., O. Ashenberg, J. Pyrdol, A.M. Luoma, O. Rozenblatt-Rosen, M. Hofree, E. Christian, L. Ferrari de Andrade, R.E. Tay, L. Teyton, A. Regev, S.K. Dougan, and K.W. Wucherpfennig. 2018. Rapid CLIP dissociation from MHC II promotes an unusual antigen presentation pathway in autoimmunity. *J Exp Med* 215:2617-2635.

Iversen, O.J., H. Lysvand, and L. Hagen. 2011. The autoantigen Pso p27: a post-translational modification of SCCA molecules. *Autoimmunity* 44:229-234.

Jamison, B.L., T. Neef, A. Goodspeed, B. Bradley, R.L. Baker, S.D. Miller, and K. Haskins. 2019. Nanoparticles Containing an Insulin-ChgA Hybrid Peptide Protect from Transfer of Autoimmune Diabetes by Shifting the Balance between Effector T Cells and Regulatory T Cells. *J Immunol* 203:48-57.

Jensen, K.K., M. Andreatta, P. Marcatili, S. Buus, J.A. Greenbaum, Z. Yan, A. Sette, B. Peters, and M. Nielsen. 2018. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154:394-406.

Jia, Y., Z. Xie, and H. Li. 2016. Intergenically Spliced Chimeric RNAs in Cancer. *Trends Cancer* 2:475-484.

Jiang, W., W.J. Swiggard, C. Heufler, M. Peng, A. Mirza, R.M. Steinman, and M.C. Nussenzweig. 1995. The receptor DEC-205 expressed by dendritic cells and thymic epithelial cells is involved in antigen processing. *Nature* 375:151-155.

Jin, N., Y. Wang, F. Crawford, J. White, P. Marrack, S. Dai, and J.W. Kappler. 2015. N-terminal additions to the WE14 peptide of chromogranin A create strong autoantigen agonists in type 1 diabetes. *Proc Natl Acad Sci U S A* 112:13318-13323.

Kakiuchi, T., R. Chesnut, and H. Grey. 1983. B cells as antigen-presenting cells: the requirement for B cell activation. *J Immunol* 131:109-114.

Kamphorst, A.O., P. Guermonprez, D. Dudziak, and M.C. Nussenzweig. 2010. Route of antigen uptake differentially impacts presentation by dendritic cells and activated monocytes. *J Immunol* 185:3426-3435.

Kerner, W., J. Bruckel, and A. German Diabetes. 2014. Definition, classification and diagnosis of diabetes mellitus. *Exp Clin Endocrinol Diabetes* 122:384-386.

Kieback, E., E. Hilgenberg, U. Stervbo, V. Lampropoulou, P. Shen, M. Bunse, Y. Jaimes, P. Boudinot, A. Radbruch, U. Klemm, A.A. Kuhl, R. Liblau, N. Hoevelmeyer, S.M. Anderton, W. Uckert, and S. Fillatreau. 2016. Thymus-Derived Regulatory T Cells Are Positively Selected on Natural Self-Antigen through Cognate Interactions of High Functional Avidity. *Immunity* 44:1114-1126.

Klein, L., B. Kyewski, P.M. Allen, and K.A. Hogquist. 2014. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol* 14:377-391.

Knight, R.R., D. Kronenberg, M. Zhao, G.C. Huang, M. Eichmann, A. Bulek, L. Wooldridge, D.K. Cole, A.K. Sewell, M. Peakman, and A. Skowera. 2013. Human beta-cell killing by autoreactive preproinsulin-specific CD8 T cells is predominantly granule-mediated with the potency dependent upon T-cell receptor avidity. *Diabetes* 62:205-213.

Koning, F., R. Thomas, J. Rossjohn, and R.E. Toes. 2015. Coeliac disease and rheumatoid arthritis: similar mechanisms, different antigens. *Nat Rev Rheumatol* 11:450-461.

Kracht, M.J., M. van Lummel, T. Nikolic, A.M. Joosten, S. Laban, A.R. van der Slik, P.A. van Veelen, F. Carlotti, E.J. de Koning, R.C. Hoeben, A. Zaldumbide, and B.O. Roep. 2017. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med* 23:501-507.

Kreiter, S., M. Vormehr, N. van de Roemer, M. Diken, M. Lower, J. Diekmann, S. Boegel, B. Schrors, F. Vascotto, J.C. Castle, A.D. Tadmor, S.P. Schoenberger, C. Huber, O. Tureci, and U. Sahin. 2015. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 520:692-696.

Krischer, J.P., D.D. Cuthbertson, L. Yu, T. Orban, N. Maclaren, R. Jackson, W.E. Winter, D.A. Schatz, J.P. Palmer, and G.S. Eisenbarth. 2003. Screening strategies for the identification of multiple antibody-positive relatives of individuals with type 1 diabetes. *J Clin Endocrinol Metab* 88:103-108.

Kronenberg, D., R.R. Knight, M. Estorninho, R.J. Ellis, M.G. Kester, A. de Ru, M. Eichmann, G.C. Huang, J. Powrie, C.M. Dayan, A. Skowera, P.A. van Veelen, and M. Peakman. 2012. Circulating preproinsulin signal peptide-specific CD8 T cells restricted by the susceptibility molecule HLA-A24 are expanded at onset of type 1 diabetes and kill beta-cells. *Diabetes* 61:1752-1759.

Laitinen, O.H., H. Honkanen, O. Pakkanen, S. Oikarinen, M.M. Hankaniemi, H. Huhtala, T. Ruokoranta, V. Lecouturier, P. Andre, R. Harju, S.M. Virtanen, J. Lehtonen, J.W. Almond, T. Simell, O. Simell, J. Ilonen, R. Veijola, M. Knip, and H. Hyoty. 2014. Coxsackievirus B1 is associated with induction of beta-cell autoimmunity that portends type 1 diabetes. *Diabetes* 63:446-455.

Larsson, P.G., T. Lakshmikanth, E. Svedin, C. King, and M. Flodstrom-Tullberg. 2013. Previous maternal infection protects offspring from enterovirus infection and prevents experimental diabetes development in mice. *Diabetologia* 56:867-874.

Lee, K.H., K.W. Wucherpfennig, and D.C. Wiley. 2001. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat Immunol* 2:501-507.

Leung, C.S., M.A. Maurer, S. Meixlsperger, A. Lippmann, C. Cheong, J. Zuo, T.A. Haigh, G.S. Taylor, and C. Munz. 2013. Robust T-cell stimulation by Epstein-Barr virus-transformed B cells after antigen targeting to DEC-205. *Blood* 121:1584-1594.

Li, Y., L. Zhou, Y. Li, J. Zhang, B. Guo, G. Meng, X. Chen, Q. Zheng, L. Zhang, M. Zhang, and L. Wang. 2015. Identification of autoreactive CD8+ T cell responses targeting chromogranin A in humanized NOD mice and type 1 diabetes patients. *Clin Immunol* 159:63-71.

Liang, S.C., X.Y. Tan, D.P. Luxenberg, R. Karim, K. Dunussi-Joannopoulos, M. Collins, and L.A. Fouser. 2006. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *J Exp Med* 203:2271-2279.

Liepe, J., F. Marino, J. Sidney, A. Jeko, D.E. Bunting, A. Sette, P.M. Kloetzel, M.P. Stumpf, A.J. Heck, and M. Mishto. 2016. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* 354:354-358.

Liepe, J., M. Mishto, K. Textoris-Taube, K. Janek, C. Keller, P. Henklein, P.M. Kloetzel, and A. Zaikin. 2010. The 20S proteasome splicing activity discovered by SpliceMet. *PLoS Comput Biol* 6:e1000830.

Liepe, J., J. Sidney, F.K.M. Lorenz, A. Sette, and M. Mishto. 2019. Mapping the MHC Class I-Spliced Immunopeptidome of Cancer Cells. *Cancer Immunol Res* 7:62-76.

Lim, J.P., and P.A. Gleeson. 2011. Macropinocytosis: an endocytic pathway for internalising large gulps. *Immunol Cell Biol* 89:836-843.

Lindley, S., C.M. Dayan, A. Bishop, B.O. Roep, M. Peakman, and T.I. Tree. 2005. Defective suppressor function in CD4(+)CD25(+) T-cells from patients with type 1 diabetes. *Diabetes* 54:92-99.

Lippolis, J.D., F.M. White, J.A. Marto, C.J. Luckey, T.N. Bullock, J. Shabanowitz, D.F. Hunt, and V.H. Engelhard. 2002. Analysis of MHC class II antigen processing by quantitation of peptides that constitute nested sets. *J Immunol* 169:5089-5097.

Lovitch, S.B., and E.R. Unanue. 2005. Conformational isomers of a peptide-class II major histocompatibility complex. *Immunol Rev* 207:293-313.

Maahs, D.M., N.A. West, J.M. Lawrence, and E.J. Mayer-Davis. 2010. Epidemiology of type 1 diabetes. *Endocrinol Metab Clin North Am* 39:481-497.

Madden, D.R. 1995. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol* 13:587-622.

Mahnke, K., M. Guo, S. Lee, H. Sepulveda, S.L. Swain, M. Nussenzweig, and R.M. Steinman. 2000. The dendritic cell receptor for endocytosis, DEC-205, can recycle and enhance antigen presentation via major histocompatibility complex class II-positive lysosomal compartments. *J Cell Biol* 151:673-684.

Mallone, R., E. Martinuzzi, P. Blancou, G. Novelli, G. Afonso, M. Dolz, G. Bruno, L. Chaillous, L. Chatenoud, J.M. Bach, and P. van Endert. 2007. CD8+ T-cell responses identify beta-cell autoimmunity in human type 1 diabetes. *Diabetes* 56:613-621.

Mannering, S.I., L.C. Harrison, N.A. Williamson, J.S. Morris, D.J. Thearle, K.P. Jensen, T.W. Kay, J. Rossjohn, B.A. Falk, G.T. Nepom, and A.W. Purcell. 2005. The insulin A-chain epitope recognized by human T cells is posttranslationally modified. *J Exp Med* 202:1191-1197.

Manoury, B., E.W. Hewitt, N. Morrice, P.M. Dando, A.J. Barrett, and C. Watts. 1998. An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature* 396:695-699.

Mason, D.M., S. Friedensohn, C.R. Weber, C. Jordi, B. Wagner, S. Meng, P. Gainza, B.E. Correia, and S.T. Reddy. 2019. Deep learning enables therapeutic antibody optimization in mammalian cells by deciphering high-dimensional protein sequence space. *bioRxiv* 617860:

Masson-Bessiere, C., M. Sebbag, E. Girbal-Neuhauser, L. Nogueira, C. Vincent, T. Senshu, and G. Serre. 2001. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. *J Immunol* 166:4177-4184.

McGinty, J.W., I.T. Chow, C. Greenbaum, J. Odegard, W.W. Kwok, and E.A. James. 2014. Recognition of posttranslationally modified GAD65 epitopes in subjects with type 1 diabetes. *Diabetes* 63:3033-3040.

Michaux, A., P. Larrieu, V. Stroobant, J.F. Fonteneau, F. Jotereau, B.J. Van den Eynde, A. Moreau-Aubry, and N. Vigneron. 2014. A spliced antigenic peptide comprising a single spliced amino acid is produced in the proteasome by reverse splicing of a longer peptide fragment followed by trimming. *J Immunol* 192:1962-1971.

Michels, A.W., L.G. Landry, K.A. McDaniel, L. Yu, M. Campbell-Thompson, W.W. Kwok, K.L. Jones, P.A. Gottlieb, J.W. Kappler, Q. Tang, B.O. Roep, M.A. Atkinson, C.E. Mathews, and M. Nakayama. 2017. Islet-Derived CD4 T Cells Targeting Proinsulin in Human Autoimmune Diabetes. *Diabetes* 66:722-734.

Mishto, M., A. Goede, K.T. Taube, C. Keller, K. Janek, P. Henklein, A. Niewienda, A. Kloss, S. Gohlke, B. Dahlmann, C. Enenkel, and P.M. Kloetzel. 2012. Driving forces of proteasome-catalyzed peptide splicing in yeast and humans. *Mol Cell Proteomics* 11:1008-1023.

Mishto, M., and J. Liepe. 2017. Post-Translational Peptide Splicing and T Cell Responses. *Trends Immunol* 38:904-915.

Mishto, M., A. Mansurkhodzhaev, G. Ying, A. Bitra, R.A. Cordfunke, S. Henze, D. Paul, J. Sidney, H. Urlaub, J. Neefjes, A. Sette, D.M. Zajonc, and J. Liepe. 2019. An in silico-in vitro Pipeline Identifying an HLA-A(*)02:01(+) KRAS G12V(+) Spliced Epitope Candidate for a Broad Tumor-Immune Response in Cancer Patients. *Front Immunol* 10:2572.

Mitelman, F., B. Johansson, and F. Mertens. 2007. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* 7:233-245.

Miyadera, H., J. Ohashi, A. Lernmark, T. Kitamura, and K. Tokunaga. 2015. Cell-surface MHC density profiling reveals instability of autoimmunity-associated HLA. *J Clin Invest* 125:275-291.

Mohan, J.F., M.G. Levisetti, B. Calderon, J.W. Herzog, S.J. Petzold, and E.R. Unanue. 2010. Unique autoreactive T cells recognize insulin peptides generated within the islets of Langerhans in autoimmune diabetes. *Nat Immunol* 11:350-354.

Mohan, J.F., S.J. Petzold, and E.R. Unanue. 2011. Register shifting of an insulin peptide-MHC complex allows diabetogenic T cells to escape thymic deletion. *J Exp Med* 208:2375-2383.

Molberg, O., S.N. McAdam, R. Korner, H. Quarsten, C. Kristiansen, L. Madsen, L. Fugger, H. Scott, O. Noren, P. Roepstorff, K.E. Lundin, H. Sjostrom, and L.M. Sollid. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med* 4:713-717.

Moss, C.X., T.I. Tree, and C. Watts. 2007. Reconstruction of a pathway of antigen processing and class II MHC peptide capture. *EMBO J* 26:2137-2147.

Moustakas, A.K., Y. van de Wal, J. Routsias, Y.M. Kooy, P. van Veelen, J.W. Drijfhout, F. Koning, and G.K. Papadopoulos. 2000. Structure of celiac disease-associated HLA-DQ8 and non-associated HLA-DQ9 alleles in complex with two disease-specific epitopes. *Int Immunol* 12:1157-1166.

Muller, B., M. Anders, H. Akiyama, S. Welsch, B. Glass, K. Nikovics, F. Clavel, H.M. Tervo, O.T. Keppler, and H.G. Krausslich. 2009. HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. *J Biol Chem* 284:29692-29703.

Nakayama, M., K. McDaniel, L. Fitzgerald-Miller, C. Kiekhaefer, J.K. Snell-Bergeon, H.W. Davidson, M. Rewers, L. Yu, P. Gottlieb, J.W. Kappler, and A. Michels. 2015. Regulatory vs. inflammatory cytokine T-cell responses to mutated insulin peptides in healthy and type 1 diabetic subjects. *Proc Natl Acad Sci U S A* 112:4429-4434.

Neefjes, J., M.L. Jongsma, P. Paul, and O. Bakke. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* 11:823-836.

Nejentsev, S., J.M. Howson, N.M. Walker, J. Szeszko, S.F. Field, H.E. Stevens, P. Reynolds, M. Hardy, E. King, J. Masters, J. Hulme, L.M. Maier, D. Smyth, R. Bailey, J.D. Cooper, G. Ribas, R.D. Campbell, D.G. Clayton, J.A. Todd, and C. Wellcome Trust Case Control. 2007. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature* 450:887-892.

Nepom, G.T., and W.W. Kwok. 1998. Molecular basis for HLA-DQ associations with IDDM. *Diabetes* 47:1177-1184.

Neugebauer, K.M., J.T. Merrill, M.H. Wener, R.G. Lahita, and M.B. Roth. 2000. SR proteins are autoantigens in patients with systemic lupus erythematosus. Importance of phosphoepitopes. *Arthritis Rheum* 43:1768-1778.

Nokoff, N., and M. Rewers. 2013. Pathogenesis of type 1 diabetes: lessons from natural history studies of high-risk individuals. *Ann N Y Acad Sci* 1281:1-15.

O'Garra, A., P.L. Vieira, P. Vieira, and A.E. Goldfeld. 2004. IL-10-producing and naturally occurring CD4+ Tregs: limiting collateral damage. *J Clin Invest* 114:1372-1378.

Oikarinen, S., S. Tauriainen, D. Hober, B. Lucas, A. Vazeou, A. Sioofy-Khojine, E. Bozas, P. Muir, H. Honkanen, J. Ilonen, M. Knip, P. Keskinen, M.T. Saha, H. Huhtala, G. Stanway, C. Bartsocas, J. Ludvigsson, K. Taylor, H. Hyoty, and G. VirDiab Study. 2014. Virus antibody survey in different European populations indicates risk association between coxsackievirus B1 and type 1 diabetes. *Diabetes* 63:655-662.

Ooi, J.D., J. Petersen, Y.H. Tan, M. Huynh, Z.J. Willett, S.H. Ramarathinam, P.J. Eggenhuizen, K.L. Loh, K.A. Watson, P.Y. Gan, M.A. Alikhan, N.L. Dudek, A. Handel, B.G. Hudson, L. Fugger, D.A. Power, S.G. Holt, P.T. Coates, J.W. Gregersen, A.W. Purcell, S.R. Holdsworth, N.L. La Gruta, H.H. Reid, J. Rossjohn, and A.R. Kitching. 2017. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. *Nature* 545:243-247.

Orr, D., T. Chen, A.H. Johnsen, R. Chalk, K.D. Buchanan, J.M.S. P, i. Rao, and C. Shaw. 2002. The spectrum of endogenous human chromogranin A-derived peptides identified using a modified proteomic strategy. *Proteomics* 2:1586-1600.

Paes, W., G. Leonov, T. Partridge, T. Chikata, H. Murakoshi, A. Frangou, S. Brackenridge, A. Nicastri, A.G. Smith, G.H. Learn, Y. Li, R. Parker, S. Oka, P. Pellegrino, I. Williams, B.F. Haynes, A.J. McMichael, G.M. Shaw, B.H. Hahn, M. Takiguchi, N. Ternette, and P. Borrow. 2019. Contribution of proteasome-catalyzed peptide cis-splicing to viral targeting by CD8(+) T cells in HIV-1 infection. *Proc Natl Acad Sci U S A* 116:24748-24759.

Panagiotopoulos, C., H. Qin, R. Tan, and C.B. Verchere. 2003. Identification of a beta-cell-specific HLA class I restricted epitope in type 1 diabetes. *Diabetes* 52:2647-2651.

Patterson, C.C., G. Dahlquist, G. Soltesz, A. Green, E.A.S.G. Europe, and Diabetes. 2001. Is childhood-onset type I diabetes a wealth-related disease? An ecological analysis of European incidence rates. *Diabetologia* 44 Suppl 3:B9-16.

Peakman, M., E.J. Stevens, T. Lohmann, P. Narendran, J. Dromey, A. Alexander, A.J. Tomlinson, M. Trucco, J.C. Gorga, and R.M. Chicz. 1999. Naturally processed and presented epitopes of the islet cell autoantigen IA-2 eluted from HLA-DR4. *J Clin Invest* 104:1449-1457.

Perrin-Cocon, L.A., C.L. Villiers, J. Salamero, F. Gabert, and P.N. Marche. 2004. B cell receptors and complement receptors target the antigen to distinct intracellular compartments. *J Immunol* 172:3564-3572.

Phelps, R.G., A.N. Turner, and A.J. Rees. 1996. Direct Identification of Naturally Processed Autoantigen-derived Peptides Bound to HLA-DR15. *Journal of Biological Chemistry* 271:18549-18553.

Pickett, B.E., D.S. Greer, Y. Zhang, L. Stewart, L. Zhou, G. Sun, Z. Gu, S. Kumar, S. Zaremba, C.N. Larsen, W. Jen, E.B. Klem, and R.H. Scheuermann. 2012. Virus pathogen database and analysis resource (ViPR): a comprehensive bioinformatics database and analysis resource for the coronavirus research community. *Viruses* 4:3209-3226.

Platteel, A.C., M. Mishto, K. Textoris-Taube, C. Keller, J. Liepe, D.H. Busch, P.M. Kloetzel, and A.J. Sijts. 2016. CD8(+) T cells of Listeria monocytogenes-infected mice recognize both linear and spliced proteasome products. *Eur J Immunol* 46:1109-1118.

Platteel, A.C.M., J. Liepe, K. Textoris-Taube, C. Keller, P. Henklein, H.H. Schalkwijk, R. Cardoso, P.M. Kloetzel, M. Mishto, and A. Sijts. 2017a. Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8(+) T Cells during Bacterial Infection. *Cell Rep* 20:1242-1253.

Platteel, A.C.M., J. Liepe, W. van Eden, M. Mishto, and A. Sijts. 2017b. An Unexpected Major Role for Proteasome-Catalyzed Peptide Splicing in Generation of T Cell Epitopes: Is There Relevance for Vaccine Development? *Front Immunol* 8:1441.

Pociot, F. 2017. Type 1 diabetes genome-wide association studies: not to be lost in translation. *Clin Transl Immunology* 6:e162.

Pociot, F., B. Akolkar, P. Concannon, H.A. Erlich, C. Julier, G. Morahan, C.R. Nierras, J.A. Todd, S.S. Rich, and J. Nerup. 2010. Genetics of type 1 diabetes: what's next? *Diabetes* 59:1561-1571.

Pociot, F., and M.F. McDermott. 2002. Genetics of type 1 diabetes mellitus. *Genes Immun* 3:235-249.

Pugliese, A. 2017. Autoreactive T cells in type 1 diabetes. *J Clin Invest* 127:2881-2891.

Purcell, A.W., S.H. Ramarathinam, and N. Ternette. 2019a. Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nat Protoc* 14:1687-1707.

Purcell, A.W., S. Sechi, and T.P. DiLorenzo. 2019b. The Evolving Landscape of Autoantigen Discovery and Characterization in Type 1 Diabetes. *Diabetes* 68:879-886.

Rammensee, H.-G., T. Friede, and S. Stevanović. 1995. MHC ligands and peptide motifs: first listing. *Immunogenetics* 41:178-228.

Redondo, M.J., M. Rewers, L. Yu, S. Garg, C.C. Pilcher, R.B. Elliott, and G.S. Eisenbarth. 1999. Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study. *BMJ* 318:698-702.

Reed, A., E. Collins, L. Shock, D. Klapper, and J. Frelinger. 1997. Diminished class II-associated Ii peptide binding to the juvenile dermatomyositis HLA-DQ alpha 1*0501/DQ beta 1*0301 molecule. *J Immunol* 159:6260-6265.

Riley, W.J., N.K. Maclaren, J. Krischer, R.P. Spillar, J.H. Silverstein, D.A. Schatz, S. Schwartz, J. Malone, S. Shah, C. Vadheim, and et al. 1990. A prospective study of the development of diabetes in relatives of patients with insulin-dependent diabetes. *N Engl J Med* 323:1167-1172.

Roche, P.A., and K. Furuta. 2015. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol* 15:203-216.

Roep, B.O. 2003. The role of T-cells in the pathogenesis of Type 1 diabetes: from cause to cure. *Diabetologia* 46:305-321.

Roep, B.O., and M. Peakman. 2012. Antigen targets of type 1 diabetes autoimmunity. *Cold Spring Harb Perspect Med* 2:a007781.

Rogers, M.A.M., C. Kim, T. Banerjee, and J.M. Lee. 2017. Fluctuations in the incidence of type 1 diabetes in the United States from 2001 to 2015: a longitudinal study. *BMC Med* 15:199.

Roncarolo, M.G., and M. Battaglia. 2007. Regulatory T-cell immunotherapy for tolerance to self antigens and alloantigens in humans. *Nat Rev Immunol* 7:585-598.

Ronningen, K.S., N. Keiding, A. Green, E.A.S.G. Europe, and Diabetes. 2001. Correlations between the incidence of childhood-onset type I diabetes in Europe and HLA genotypes. *Diabetologia* 44 Suppl 3:B51-59.

Rorsman, P., and E. Renstrom. 2003. Insulin granule dynamics in pancreatic beta cells. *Diabetologia* 46:1029-1045.

Sandberg, M., and L.A. Borg. 2006. Intracellular degradation of insulin and crinophagy are maintained by nitric oxide and cyclo-oxygenase 2 activity in isolated pancreatic islets. *Biol Cell* 98:307-315.

Sant'Angelo, D.B., E. Robinson, C.A. Janeway, Jr., and L.K. Denzin. 2002. Recognition of core and flanking amino acids of MHC class II-bound peptides by the T cell receptor. *Eur J Immunol* 32:2510-2520.

Scally, S.W., J. Petersen, S.C. Law, N.L. Dudek, H.J. Nel, K.L. Loh, L.C. Wijeyewickrema, S.B. Eckle, J. van Heemst, R.N. Pike, J. McCluskey, R.E. Toes, N.L. La Gruta, A.W. Purcell, H.H. Reid, R. Thomas, and J. Rossjohn. 2013. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J Exp Med* 210:2569-2582.

Schmid, D., M. Pypaert, and C. Munz. 2007. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* 26:79-92.

Schvartz, D., Y. Brunner, Y. Coute, M. Foti, C.B. Wollheim, and J.C. Sanchez. 2012. Improved characterization of the insulin secretory granule proteomes. *J Proteomics* 75:4620-4631.

Serr, I., and C. Daniel. 2018. Regulation of T Follicular Helper Cells in Islet Autoimmunity. *Front Immunol* 9:1729.

Shojaeian, A., and A. Mehri-Ghahfarrokhi. 2018. An overview of the Epidemiology of Type 1 Diabetes Mellitus. *Int J Metab Syndr* 2:001-004.

Skowera, A., R.J. Ellis, R. Varela-Calvino, S. Arif, G.C. Huang, C. Van-Krinks, A. Zaremba, C. Rackham, J.S. Allen, T.I. Tree, M. Zhao, C.M. Dayan, A.K. Sewell, W.W. Unger, J.W. Drijfhout, F. Ossendorp, B.O. Roep, and M. Peakman. 2008. CTLs are targeted to kill beta cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope. *J Clin Invest* 118:3390-3402.

So, M., C.M. Elso, E. Tresoldi, M. Pakusch, V. Pathiraja, J.M. Wentworth, L.C. Harrison, B. Krishnamurthy, H.E. Thomas, C. Rodda, F.J. Cameron, J. McMahon, T.W.H. Kay, and S.I. Mannering. 2018. Proinsulin C-peptide is an autoantigen in people with type 1 diabetes. *Proc Natl Acad Sci U S A* 115:10732-10737.

Sobota, J.A., N. Back, B.A. Eipper, and R.E. Mains. 2009. Inhibitors of the V0 subunit of the vacuolar H+-ATPase prevent segregation of lysosomal- and secretory-pathway proteins. *J Cell Sci* 122:3542-3553.

Sollid, L.M., and B. Jabri. 2011. Celiac disease and transglutaminase 2: a model for posttranslational modification of antigens and HLA association in the pathogenesis of autoimmune disorders. *Curr Opin Immunol* 23:732-738.

Soltesz, G., C.C. Patterson, and G. Dahlquist. 2007. Worldwide childhood type 1 diabetes incidence ? what can we learn from epidemiology? *Pediatric Diabetes* 8:6-14.

Spence, A., W. Purtha, J. Tam, S. Dong, Y. Kim, C.H. Ju, T. Sterling, M. Nakayama, W.H. Robinson, J.A. Bluestone, M.S. Anderson, and Q. Tang. 2018. Revealing the specificity of regulatory T cells in murine autoimmune diabetes. *Proc Natl Acad Sci U S A* 115:5265-5270.

Stadinski, B.D., T. Delong, N. Reisdorph, R. Reisdorph, R.L. Powell, M. Armstrong, J.D. Piganelli, G. Barbour, B. Bradley, F. Crawford, P. Marrack, S.K. Mahata, J.W. Kappler, and K. Haskins. 2010a. Chromogranin A is an autoantigen in type 1 diabetes. *Nat Immunol* 11:225-231.

Stadinski, B.D., L. Zhang, F. Crawford, P. Marrack, G.S. Eisenbarth, and J.W. Kappler. 2010b. Diabetogenic T cells recognize insulin bound to IAg7 in an unexpected, weakly binding register. *Proc Natl Acad Sci U S A* 107:10978-10983.

Standifer, N.E., Q. Ouyang, C. Panagiotopoulos, C.B. Verchere, R. Tan, C.J. Greenbaum, C. Pihoker, and G.T. Nepom. 2006. Identification of Novel HLA-A*0201-restricted epitopes in recent-onset type 1 diabetic subjects and antibody-positive relatives. *Diabetes* 55:3061-3067.

Starck, S.R., and N. Shastri. 2011. Non-conventional sources of peptides presented by MHC class I. *Cell Mol Life Sci* 68:1471-1479.

Stepniak, D., M. Wiesner, A.H. de Ru, A.K. Moustakas, J.W. Drijfhout, G.K. Papadopoulos, P.A. van Veelen, and F. Koning. 2008. Large-scale characterization of natural ligands explains the unique gluten-binding properties of HLA-DQ2. *J Immunol* 180:3268-3278.

Stevens, E.J., and M. Peakman. 1998. Enhanced T cell proliferation and increased responder frequency following delivery of antigen to the antigen-presenting cell; B cell dependency and use in detection of autoreactive T cells. *J Immunol Methods* 215:59-70.

Strollo, R., C. Vinci, N. Napoli, P. Pozzilli, J. Ludvigsson, and A. Nissim. 2017. Antibodies to post-translationally modified insulin as a novel biomarker for prediction of type 1 diabetes in children. *Diabetologia* 60:1467-1474.

Stuart, L.M., and R.A. Ezekowitz. 2005. Phagocytosis: elegant complexity. *Immunity* 22:539-550.

Suckale, J., and M. Solimena. 2010. The insulin secretory granule as a signaling hub. *Trends Endocrinol Metab* 21:599-609.

Suri, A., J.J. Walters, M.L. Gross, and E.R. Unanue. 2005. Natural peptides selected by diabetogenic DQ8 and murine I-A(g7) molecules show common sequence specificity. *J Clin Invest* 115:2268-2276.

Tachibana, H., K. Jiyoun, K. Taniguchi, Y. Ushio, K. Teruya, K. Osada, Y. Inoue, S. Shirahata, and H. Murakami. 1996. Modified antigen-binding of human antibodies with glycosylation variations of the light chains produced in sugar-limited human hybridoma cultures. *In Vitro Cell Dev Biol Anim* 32:178-183.

Thunander, M., C. Petersson, K. Jonzon, J. Fornander, B. Ossiansson, C. Torn, S. Edvardsson, and M. Landin-Olsson. 2008. Incidence of type 1 and type 2 diabetes in adults and children in Kronoberg, Sweden. *Diabetes Res Clin Pract* 82:247-255.

Todd, J.A., J.I. Bell, and H.O. McDevitt. 1987. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* 329:599-604.

Trouw, L.A., T. Rispens, and R.E.M. Toes. 2017. Beyond citrullination: other post-translational protein modifications in rheumatoid arthritis. *Nat Rev Rheumatol* 13:331-339.

UniProt Consortium. 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 40:D71-75.

Valle, A., G.M. Giamporcaro, M. Scavini, A. Stabilini, P. Grogan, E. Bianconi, G. Sebastiani, M. Masini, N. Maugeri, L. Porretti, R. Bonfanti, F. Meschi, M. De Pellegrin, A. Lesma, S. Rossini, L. Piemonti, P. Marchetti, F. Dotta, E. Bosi, and M. Battaglia. 2013. Reduction of circulating neutrophils precedes and accompanies type 1 diabetes. *Diabetes* 62:2072-2077.

van Belle, T.L., K.T. Coppieters, and M.G. von Herrath. 2011. Type 1 diabetes: etiology, immunology, and therapeutic strategies. *Physiol Rev* 91:79-118.

van de Wal, Y., Y. Kooy, P. van Veelen, S. Pena, L. Mearin, G. Papadopoulos, and F. Koning. 1998. Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J Immunol* 161:1585-1588.

van Lummel, M., D.T.P. Buis, C. Ringeling, A.H. de Ru, J. Pool, G.K. Papadopoulos, P.A. van Veelen, H. Reijonen, J.W. Drijfhout, and B.O. Roep. 2019. Epitope Stealing as a Mechanism of Dominant Protection by HLA-DQ6 in Type 1 Diabetes. *Diabetes* 68:787-795.

van Lummel, M., G. Duinkerken, P.A. van Veelen, A. de Ru, R. Cordfunke, A. Zaldumbide, I. Gomez-Tourino, S. Arif, M. Peakman, J.W. Drijfhout, and B.O. Roep. 2014. Posttranslational modification of HLA-DQ binding islet autoantigens in type 1 diabetes. *Diabetes* 63:237-247.

van Lummel, M., P.A. van Veelen, A.H. de Ru, G.M. Janssen, J. Pool, S. Laban, A.M. Joosten, T. Nikolic, J.W. Drijfhout, M.L. Mearin, H.J. Aanstoot, M. Peakman, and B.O. Roep. 2016a. Dendritic Cells Guide Islet Autoimmunity through a Restricted and Uniquely Processed Peptidome Presented by High-Risk HLA-DR. *J Immunol* 196:3253-3263.

van Lummel, M., P.A. van Veelen, A.H. de Ru, J. Pool, T. Nikolic, S. Laban, A. Joosten, J.W. Drijfhout, I. Gomez-Tourino, S. Arif, H.J. Aanstoot, M. Peakman, and B.O. Roep. 2016b. Discovery of a Selective Islet Peptidome Presented by the Highest-Risk HLA-DQ8trans Molecule. *Diabetes* 65:732-741.

Vermeulen, I., I. Weets, M. Asanghanwa, J. Ruige, L. Van Gaal, C. Mathieu, B. Keymeulen, V. Lampasona, J.M. Wenzlau, J.C. Hutton, D.G. Pipeleers, F.K. Gorus, and R. Belgian Diabetes. 2011. Contribution of antibodies against IA-2beta and zinc transporter 8 to classification of diabetes diagnosed under 40 years of age. *Diabetes Care* 34:1760-1765.

Verspurten, J., K. Gevaert, W. Declercq, and P. Vandenabeele. 2009. SitePredicting the cleavage of proteinase substrates. *Trends Biochem Sci* 34:319-323.

Vigneron, N., V. Ferrari, V. Stroobant, J. Abi Habib, and B.J. Van den Eynde. 2017. Peptide splicing by the proteasome. *J Biol Chem* 292:21170-21179.

Vigneron, N., V. Stroobant, J. Chapiro, A. Ooms, G. Degiovanni, S. Morel, P. van der Bruggen, T. Boon, and B.J. Van den Eynde. 2004. An antigenic peptide produced by peptide splicing in the proteasome. *Science* 304:587-590.

Vigneron, N., V. Stroobant, V. Ferrari, J. Abi Habib, and B.J. Van den Eynde. 2019. Production of spliced peptides by the proteasome. *Mol Immunol* 113:93-102.

Wahlstrom, J., J. Dengjel, B. Persson, H. Duyar, H.G. Rammensee, S. Stevanovic, A. Eklund, R. Weissert, and J. Grunewald. 2007. Identification of HLA-DR-bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis. *J Clin Invest* 117:3576-3582.

Walker, L.S., and A.K. Abbas. 2002. The enemy within: keeping self-reactive T cells at bay in the periphery. *Nat Rev Immunol* 2:11-19.

Wallberg, M., and A. Cooke. 2013. Immune mechanisms in type 1 diabetes. *Trends Immunol* 34:583-591.

Walter, U., T. Toepfer, K.E. Dittmar, K. Kretschmer, J. Lauber, S. Weiss, G. Servos, O. Lechner, W.A. Scherbaum, S.R. Bornstein, H. Von Boehmer, and J. Buer. 2003. Pancreatic NOD beta cells express MHC class II protein and the frequency of I-A(g7) mRNA-expressing beta cells strongly increases during progression to autoimmune diabetes. *Diabetologia* 46:1106-1114.

Wan, X., A.N. Vomund, O.J. Peterson, A.V. Chervonsky, C.F. Lichti, and E.R. Unanue. 2020. The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. *Nat Immunol* 21:455-463.

Wan, X., B.H. Zinselmeyer, P.N. Zakharov, A.N. Vomund, R. Taniguchi, L. Santambrogio, M.S. Anderson, C.F. Lichti, and E.R. Unanue. 2018. Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* 560:107-111.

Wang, N., E. Weber, and J.S. Blum. 2009. Diminished intracellular invariant chain expression after vaccinia virus infection. *J Immunol* 183:1542-1550.

Wang, Y., T. Sosinowski, A. Novikov, F. Crawford, J. White, N. Jin, Z. Liu, J. Zou, D. Neau, H.W. Davidson, M. Nakayama, W.W. Kwok, L. Gapin, P. Marrack, J.W. Kappler, and S. Dai. 2019. How C-terminal additions to insulin B-chain fragments create superagonists for T cells in mouse and human type 1 diabetes. *Sci Immunol* 4:

Warren, E.H., N.J. Vigneron, M.A. Gavin, P.G. Coulie, V. Stroobant, A. Dalet, S.S. Tykodi, S.M. Xuereb, J.K. Mito, S.R. Riddell, and B.J. Van den Eynde. 2006. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* 313:1444-1447.

Warren, R.L., J.D. Freeman, T. Zeng, G. Choe, S. Munro, R. Moore, J.R. Webb, and R.A. Holt. 2011. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* 21:790-797.

Wen, X., J. Yang, E. James, I.T. Chow, H. Reijonen, and W.W. Kwok. 2020. Increased islet antigen-specific regulatory and effector CD4(+) T cells in healthy individuals with the type 1 diabetes-protective haplotype. *Sci Immunol* 5:

Wicker, L.S., S.L. Chen, G.T. Nepom, J.F. Elliott, D.C. Freed, A. Bansal, S. Zheng, A. Herman, A. Lernmark, D.M. Zaller, L.B. Peterson, J.B. Rothbard, R. Cummings, and P.J. Whiteley. 1996. Naturally processed T cell epitopes from human glutamic acid decarboxylase identified using mice transgenic for the type 1 diabetes-associated human MHC class II allele, DRB1*0401. *J Clin Invest* 98:2597-2603.

Wiesner, M., D. Stepniak, A.H. de Ru, A.K. Moustakis, J.W. Drijfhout, G.K. Papadopoulos, P.A. van Veelen, and F. Koning. 2008. Dominance of an alternative CLIP sequence in the celiac disease associated HLA-DQ2 molecule. *Immunogenetics* 60:551-555.

Wiles, T.A., T. Delong, R.L. Baker, B. Bradley, G. Barbour, R.L. Powell, N. Reisdorph, and K. Haskins. 2017. An insulin-IAPP hybrid peptide is an endogenous antigen for CD4 T cells in the non-obese diabetic mouse. *J Autoimmun* 78:11-18.

Wiles, T.A., R. Powell, R. Michel, K.S. Beard, A. Hohenstein, B. Bradley, N. Reisdorph, K. Haskins, and T. Delong. 2019. Identification of Hybrid Insulin Peptides (HIPs) in Mouse and Human Islets by Mass Spectrometry. *J Proteome Res* 18:814-825.

Wong, F.S., L. Wen, M. Tang, M. Ramanathan, I. Visintin, J. Daugherty, L.G. Hannum, C.A. Janeway, Jr., and M.J. Shlomchik. 2004. Investigation of the role of B-cells in type 1 diabetes in the NOD mouse. *Diabetes* 53:2581-2587.

Wooldridge, L., J. Ekeruche-Makinde, H.A. van den Berg, A. Skowera, J.J. Miles, M.P. Tan, G. Dolton, M. Clement, S. Llewellyn-Lacey, D.A. Price, M. Peakman, and A.K. Sewell. 2012. A single autoimmune T cell receptor recognizes more than a million different peptides. *J Biol Chem* 287:1168-1177.

Xia, Y., Z. Xie, G. Huang, and Z. Zhou. 2019. Incidence and trend of type 1 diabetes and the underlying environmental determinants. *Diabetes Metab Res Rev* 35:e3075.

Yamaguchi, K., M. Uechi, Y. Katakura, T. Oda, and M. Ishiguro. 2004. Mitogenic properties of pokeweed lectin-D isoforms on human peripheral blood lymphocytes: non-mitogen PL-D1 and mitogen PL-D2. *Biosci Biotechnol Biochem* 68:1591-1593.

Yang, J., I.T. Chow, T. Sosinowski, N. Torres-Chinn, C.J. Greenbaum, E.A. James, J.W. Kappler, H.W. Davidson, and W.W. Kwok. 2014. Autoreactive T cells specific for insulin B:11-23 recognize a low-affinity peptide register in human subjects with autoimmune diabetes. *Proc Natl Acad Sci U S A* 111:14840-14845.

Yang, J., N.A. Danke, D. Berger, S. Reichstetter, H. Reijonen, C. Greenbaum, C. Pihoker, E.A. James, and W.W. Kwok. 2006. Islet-specific glucose-6-phosphatase catalytic subunit-related protein-reactive CD4+ T cells in human subjects. *J Immunol* 176:2781-2789.

Yarchoan, M., B.A. Johnson, 3rd, E.R. Lutz, D.A. Laheru, and E.M. Jaffee. 2017. Targeting neoantigens to augment antitumour immunity. *Nat Rev Cancer* 17:209-222.

Yewdell, J.W. 2011. DRiPs solidify: progress in understanding endogenous MHC class I antigen processing. *Trends Immunol* 32:548-558.

Yoon, J.W., M. Austin, T. Onodera, and A.L. Notkins. 1979. Isolation of a virus from the pancreas of a child with diabetic ketoacidosis. *N Engl J Med* 300:1173-1179.

Zamvil, S.S., D.J. Mitchell, A.C. Moore, K. Kitamura, L. Steinman, and J.B. Rothbard. 1986. T-cell epitope of the autoantigen myelin basic protein that induces encephalomyelitis. *Nature* 324:258-260.

Zhao, Y., N.A. Scott, H.S. Quah, B. Krishnamurthy, F. Bond, T. Loudovaris, S.I. Mannering, T.W. Kay, and H.E. Thomas. 2015. Mouse pancreatic beta cells express MHC class II and stimulate CD4(+) T cells to proliferate. *Eur J Immunol* 45:2494-2503.

Zhou, W., S. Yang, and P.G. Wang. 2017a. Matrix effects and application of matrix effect factor. *Bioanalysis* 9:1839-1844.

Zhou, Z., E. Reyes-Vargas, H. Escobar, K.Y. Chang, A.P. Barker, A.L. Rockwood, J.C. Delgado, X. He, and P.E. Jensen. 2017b. Peptidomic analysis of type 1 diabetes associated HLA-DQ molecules and the impact of HLA-DM on peptide repertoire editing. *Eur J Immunol* 47:314-326.

Ziegler, A.G., M. Rewers, O. Simell, T. Simell, J. Lempainen, A. Steck, C. Winkler, J. Ilonen, R. Veijola, M. Knip, E. Bonifacio, and G.S. Eisenbarth. 2013. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* 309:2473-2479.